AFRL-RI-RS-TR-2013-058

OPTIMAL METHODS FOR CLASSIFICATION OF DIGITALLY MODULATED SIGNALS

MARCH 2013

INTERIM TECHNICAL REPORT

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED

STINFO COPY

AIR FORCE RESEARCH LABORATORY
INFORMATION DIRECTORATE

■ **AIR FORCE MATERIEL COMMAND** ■ **UNITED STATES AIR FORCE** ■ **ROME, NY 13441**

NOTICE AND SIGNATURE PAGE

AFRL-RI-RS-TR-2013-058 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

FOR THE DIRECTOR:

/ S /

BRENT HOLMES
Chief, Cyber Operations Branch

/ S /

WARREN H. DEBANY, JR.
Technical Advisor, Information Exploitation & Operations Division
Information Directorate

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From - To)*
MARCH 2013	INTERIM TECHNICAL REPORT	OCT 2011 – OCT 2012

4. TITLE AND SUBTITLE

OPTIMAL METHODS FOR CLASSIFICATION OF DIGITALLY MODULATED SIGNALS

5a. CONTRACT NUMBER IN-HOUSE

5b. GRANT NUMBER N/A

5c. PROGRAM ELEMENT NUMBER 63788F

6. AUTHOR(S)

Alfredo Vega Irizarry

5d. PROJECT NUMBER GCIH

5e. TASK NUMBER OM

5f. WORK UNIT NUMBER MC

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/Information Directorate
Rome Research Site/RIGB
525 Brooks Road
Rome NY 13441-4505

8. PERFORMING ORGANIZATION REPORT NUMBER

N/A

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Air Force Research Laboratory/Information Directorate
Rome Research Site/RIGB
525 Brooks Road
Rome NY 13441-4505

10. SPONSOR/MONITOR'S ACRONYM(S)

N/A

11. SPONSORING/MONITORING AGENCY REPORT NUMBER

AFRL-RI-RS-TR-2013-058

12. DISTRIBUTION AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited. PA# 88ABW-2013-1042
Date Cleared:5 MARCH 2013

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report summarizes the work done under the in-house effort titled “Optimal Methods for Modulation Classification.” The effort revisited the problem of blind demodulation for develop classification algorithms for wider set of signals types. Two methodologies were used: Likelihood Ratio Test and inference methods using Expectation Maximization. A simple case of the likelihood was developed for spread spectrum. Also a minimum divergence approach was developed for classifying MPSK and QAM. Future work will refine these two proposed concepts. Finally, this research explored MIMO codes design. The purpose was to design codes that exceeded the performance of the Full-Rate, Full Diversity Golden Codes. It was possible to find linear codes with higher spectral rate, but also higher average power. This concepts will be refined as part of future work.

15. SUBJECT TERMS

Classification, Detection, Modulation, LRT, Expectation Maximization

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			**ALFREDO VEGA IRIZARRY**
U	U	U	UU	65	19b. TELEPONE NUMBER (*Include area code*) **N/A**

Contents

Table of Figures

List of Tables

1 Abstract

This report summarizes the research done under the in-house effort titled "Optimal Methods for Modulation Classification." The purpose of this research is to revisit the problem of blind demodulation and develop classification algorithms for wide set of signals types. The effort explores two methodologies in Decision Theory: Likelihood Ratio Test and inference methods using Expectation Maximization. The focus will be mainly in types such as QAM, MPSK and BPSK Spread Spectrum. The algorithms must have medium complexity and provide some optimal classification.

In this report, we explored the challenges of constructing LRT rules for BPSK spread spectrum. Simple assumptions such as assuming statistically independent waveform coefficients proved to be inappropriate to develop such rule. An assumption of using a probability distribution based on the Total Square Correlation resulted to be effective, although it turn out to be a complicated procedure for waveforms coefficients derived from 4x4 Hadamard matrices.

The second approach used was based on Information Divergence. This approach can be adapted for multiple modulation schemes. It was effective for simple modulation schemes such as BPSK, QPSK and QAM16, but it has convergence problems for complex constellations. Future work will apply the same technique to spread spectrum signals and will try to improve the convergence.

Finally, this research explored MIMO codes design. The purpose was to design codes that exceeded the performance of the Full-Rate, Full Diversity Golden Codes. The approach consisted in searching for linear 2x2 MIMO codes that can allocate 6 symbols instead of 4. Although it was possible to find such codes; however, none of the codes that were found exceeded the performance of the Golden Code.

2 Introduction

Existing methods in modulation classification can be grouped in three categories: heuristic methods, feature based and decision theory classification.

Heuristic methods are derived from the observation of the signal in time and frequency domains. Examples of these are: the Phase Histogram classifier was proposed by Lifdtke (Lifdtke, 1984) and the well-known M^{th} Law Classifier for MPSK signals. The performance of these methods is unpredictable. From a research point of view, heuristic methods have minor importance in the area of modulation classification.

Feature-based methods rely on the heuristic selection of features, but introduce a more scientific approach by using clustering algorithms for decision making. These methods are more popular in the literature and they make use of statistical moments. (Dobre, Abdi, Bar-Ness, & Su, 2007) They use data set for the characterization of the signals under classification. The performance of these methods is known to depend on the training data. Examples of these methods are the Statistical-Moment Based Classifier presented by Soliman (Soliman & Sue, 1992) and the Sombrero Classifier developed by Army RDECOM.

Finally, the decision theoretic methods are based in Bayes Decision Theory. These methods offer superior performance because they are optimized for minimum probability of error. (Huang & Polydoros, 1995) Methods based on optimal decision theory are less abundant in the literature due to the complexity in the development of such algorithms. They require good models and some appropriate optimization criterion for deriving the optimal detector. The algorithms are designed to guarantee the best performance in low SNR. The few known methods of this kind are signal-specific and most of them apply simple MPSK and QAM models.

The purpose of this research is to extend the decision theoretic methods to broader sets of signals while trying to minimize the complexity of the algorithms. To achieve this goal, we will explore two paths: the development of likelihood approaches and the use of an entropy based method known as the Minimum Divergence Principle (MDP).

The discussion of this report is organized as follows: First, the MPSK Likelihood Ratio Test will be revisited and modified to accommodate spread spectrum signals. As an alternative approach, inference methods such as EM and MDP will be applied to MPSK/QAM and then extended to spread spectrum. Apart from decision theoretic methods, the report explores the design of MIMO codes for optimal performance. MIMO codes are considered a spatial-temporal modulation with the purpose of understanding MIMO codes and extending classification algorithms to this set of signals.

3 Modulation Classification Methods

3.1 Existing Likelihood Methods

Various classification methods using decision theory can be found in the literature. These methods targets a specific set of signal that include a combination of MPSK and QAM. Table 1 shows some of these developments. (Dobre, Abdi, Bar-Ness, & Su, 2007)

Table 1. Existing Decision Theoretic Methods

Method	Signal Types	Authors
ALRT	BPSK, QPSK	Huang, Polydoros (Kim & Polydoros, 1988)
ALRT	MPSK	Huang, Polydoros (Huang & Polydoros, 1995)
ALRT	16PSK, 16QAM	Sapiano (Sapiano & Martin, 1996)
GLRT	16PSK, 16QAM	Panagiotou (Panagioutou, Anastasoupoulos, & Polydoros, 2000)
HLRT	16PSK, 16QAM, 64QAM	Dobre (Dobre, Modulation Classification in Fading Channels using Antenna Arrays, 2004)
EM	BPSK-SS	Yao, Poor (Yao & Poor, 2000)

The list is not exhaustive, but helps to understand some of the challenges in the area of research. Average Likelihood Ratio Test (ALRT) requires taking the expectation of all unknown variables. The unknown variables are: the symbols, which are usually several hundreds or thousands; the phase offset; and the time offset. Integrating over hundreds symbols is possible if the symbols are considered statically independent, then the expectation (i.e., integration) of products becomes the product of the expectation (i.e., integration).

Some developments have tried to simplify the integration by performing a finite sum. These methods are known as Quasi-ALRT. Other methods just provide a good estimate of some unknown parameters. The methods are known as Generalized LRT. Other methods contain a mixed of methodologies. These methods are called Hybrid-LRT.

The methods in Table 1 may correspond to one of the types listed in Table 2.

Table 2. Summary of Modulation Classification Methods

Likelihood Based		**Feature Based**
Likelihood Ratio Test (LRT)	**Expectation Maximization Approaches (EM/MDP)**	**Statistical, Histogram and Spectral Features**
Goal: Classification of MPSK or QAM	**Goal:** Classification of MPSK, QAM and Spread Spectrum	**Goal:** Classification of MPSK, QAM
Approach: Derivation of a rule for classification using the signal model	**Approach:** Inference of the likelihood distribution from the correlator output	**Approach:** Generation of features using peak power and frequency information; classify signals using deviations in frequency or amplitude
Assumptions: Test a single test signal, Cases: Coherent Detection, Non-coherent Detection, or Asynchronous detection	**Assumptions:** Test a single signal, Requires time synchronization, Non-coherent Detection	**Assumptions:** Test multiple non-overlapping signals in a wide spectrum. May have multiple modulation types.

Likelihood Based		Feature Based
Likelihood Ratio Test (LRT)	**Expectation Maximization Approaches (EM/MDP)**	**Statistical, Histogram and Spectral Features**
Signal Model: $s(t) = \sum_{k=0}^{K-1} \sqrt{2E}\, b_k\, e^{j\theta_o} p(t - kT)$ $r(t) = s(t) + n(t)$ $n(t) \sim AWGN$	**Stochastic Model:** $p(\mathrm{r}\vert\mathrm{z};\vec{\theta}) = \frac{1}{(\pi N_0)^{1/2}} \exp\left(-\frac{\left\vert \mathrm{r} - \mathrm{z}e^{j\theta}\right\vert^2}{N_0}\right)$ $q(r\vert H_M) = \sum_{l=0}^{M-1} \frac{1}{M(\pi N_0)^{1/2}} exp\left(-\frac{\left\vert r - e^{\frac{j2\pi l}{M}}\right\vert^2}{N_0}\right)$	**Signal Model:** $s(n) = \sum_{m=0}^{K-1} S_k(n)\, e^{j(2\pi f_m + \theta_m)} + N(n)$ Mixture of signals is possible
Classification Criteria: Likelihood ratio test: $\Lambda(\mathrm{r}) = \frac{E_p\{\lambda(r\vert H_1)\}}{E_p\{\lambda(r\vert H_0)\}} \underset{H_0}{\overset{H_1}{\gtrless}} \eta$ Conditional Likelihood: $\lambda = e^{\frac{2}{N_0}\int_0^{NT_s} Re\{\, r(t)\cdot s^*(t;p\vert H_1)\}dt}$	**Classification Criteria:** Minimum Divergence $D_{KL}\left(q(r,z;\vec{\theta})\vert\vert\, p(r,z;\vec{\theta})\right)$ After parameter estimation: $\vec{\theta}^{opt} = \arg\max_{\vec{\theta}} D_{KL}\left(q(r,z;\vec{\theta})\vert\vert\, p(r,z;\vec{\theta})\right)$	**Classification Criteria:** Decision Tree Classification using features from histogram
Strengths: Optimal performance	**Strengths:** Does not require complex equations, extended to Spread Spectrum	**Strengths:** Easier implementation
Weaknesses: Complex decision rules or non-analytical. Limited implementation to MPSK and QAM	**Weaknesses:** Convergence problems; multiple local minima	**Weaknesses:** Susceptible to power fluctuations; Inability to differentiate narrow band signals and tones.

3.2 Likelihood Ratio Test (LRT)

One of the most relevant papers found in modulation classification is "*Likelihood Methods for MPSK Classification*". This was the first method in departing from feature-based classification by developing simple rules for classifying the set of MPSK, i.e., BPSK, QPSK, 8PSK, 16PSK and others. The method validated existing heuristic Mth power law. It showed that the developed rules were special cases of a more general decision rule. The method served as a starting point for developing classifiers of spread spectrum signals.

3.2.1 MPSK Model

We consider a baseband, finite sequence of symbols b_k with a phase offset $e^{j\theta_o}$ carried by non-overlapping normalized pulses $p(t)$. The number of symbols is K and the duration of each pulse is T. The signal $s(t)$ is a superposition of pulses modulated by the respective symbols, as shown in (1). The energy of each symbol is represented by E.

$$s(t) = \sum_{k=0}^{K-1} \sqrt{E}\, b_k\, e^{j\theta_o} p(t-kT) \quad for\ \ 0 \le t \le NT \tag{1}$$

For MPSK information symbols the symbols have amplitude of one and discrete phases which are uniformly distributed around the unit circle, (2). The M represents the number of symbols. In the case of BPSK signal, $M = 2$.

$$b_k = e^{j\theta_k}, \qquad \theta_k = \frac{2\pi k}{M}, \quad k = 0,1,\dots,M \tag{2}$$

3.2.2 Decision Theoretic Approach

The *likelihood function* is just a density function conditioned to a given hypothesis. (3) If the stochastic model is a composite model with several parameters beside the random variables, the probability density function is referred as a *conditional likelihood function*.

$$\lambda(X|H) = p(X|H) \tag{3}$$

The *likelihood ratio* is the ratio of two likelihood functions under two hypotheses.

$$\Lambda(X) = \frac{\lambda(X|H_A)}{\lambda(X|H_B)} \tag{4}$$

When we deal with signals, we talk about processes. A *process* is a random event that can be characterized with a parametric probability density function (pdf). The random variable X is replaced with all the possible indexed variables $r(t)$. If $r(t)$ is a set of Gaussian independent variables the variance is infinite:

$$\lambda(r(t)|H) = \lim_{\Delta t \to 0} \prod_{\substack{t=0 \\ t=t+\Delta t}}^{t_f} \frac{1}{\sqrt{\pi N_0}} e^{-|r(t)-s(t)|^2/N_0} \tag{5}$$

Redefining a likelihood ratio for process requires introducing the notion of orthogonal basis representation. A process $n(t)$ will be decomposed in arbitrary but deterministic basis functions $\varphi_i(t)$ as shown in (6). Using the assumption of a wide sense stationary process (7)-(8) and treating the transmitted signal as a deterministic process, it is possible to compute the likelihood ratio for a finite bandwidth process. (See Appendix A)

$$n(t) = \sum_{i=0}^{\infty} n_i \varphi_i(t) \tag{6}$$

$$E\{n(t)\} = 0 \tag{7}$$

$$E\{n(t)n(u)\} = \frac{N_0}{2} \delta(t-u) \tag{8}$$

For an AGWN process, the likelihood functions are defined according to (9). The hypothesis H is implicitly stated in the model chosen for the transmitted signal $s(t)$.

$$\lambda(r(t)|H) = e^{-\frac{(r(t)-s(t))^2}{N_0}} \tag{9}$$

Most of the complexity of (9) appears when we consider unknown parameters such as symbols b_k, phase offset θ_o (phase synchronization), symbol period, and time offset (time synchronization). In this case, the conditional likelihood needs to be averaged over all possible symbols.

3.2.3 Case 1: Averaging Over Unknown Symbols

Averaging over symbols is done under the assumption of a known distribution. In the worst case, it is assumed that the symbols are equally distributed. In the case of MPSK, we have:

$$p\left(b_k = e^{j\theta}\right) = \frac{1}{M}\,\delta\left(b_k - e^{j\theta}\right) \tag{10}$$

$$p(b_k|MPSK) = \frac{1}{M}\sum_{m=0}^{M-1}\delta\left(b_k - e^{j(\theta_m-\theta_0)}\right) \tag{11}$$

$$\lambda(r(t)|H) = \mathbb{E}_{b_0 b_1 \dots b_K}\left\{e^{-\frac{(r(t)-s(t))^2}{N_0}}\right\} \tag{12}$$

The statistical independence of each symbol allows separating the expectation of all symbols into the product of the expectations (13). This may bring some computational benefits.

$$\mathbb{E}_{b_0 b_1 \dots b_{K-1}}\left\{e^{-\frac{(r(t)-s(t))^2}{N_0}}\right\} = \prod_{i=0}^{K-1}\mathbb{E}_{b_i}\left\{e^{-\frac{(r(t)-s(t))^2}{N_0}}\right\} \tag{13}$$

3.2.4 Case 2: Averaging Over Unknown Phase

The previous equation is a case of the GLRT when the phase offset θ_0 is unknown. The generalized likelihood method uses the best estimate of θ_0 rather than averaging over all possible values of the parameter.

For MPSK case, the distribution of θ_0 is assumed to be uniform, i.e., there is no preference in the values.

$$p(\theta_0) = \frac{1}{2\pi} \tag{14}$$

$$\lambda(r(t)|H) = \mathbb{E}_{\theta_0}\left\{\mathbb{E}_{b_0 b_1 \ldots b_K}\left\{e^{-\frac{(r(t)-s(t))^2}{N_0}}\right\}\right\} \tag{15}$$

3.2.5 Case 3: Averaging Over Unknown Timing

Perhaps, the most important parameters are time offset and symbol period. Recovering the symbols would be almost impossible without any of these parameters. Yet, many literature papers make the assumptions that unknown signals are perfectly synchronized. One may argue that a preamble must be present in any communication system and thus it is possible to synchronize signals with high degree of accuracy.

The time offset synchronization is modeled by an error in the limits of the integration in the correlator as shown in (16) and (17). Equation (18) is the overall expression for the likelihood of a signal with unknown data, phase offset and time offset.

$$r_k(\epsilon) = \int_{kT}^{(k+1)T} r(t-\epsilon T)p(t)^* dt \tag{16}$$

$$r_k(\epsilon) = \int_{(k-\epsilon)T}^{(k+1-\epsilon)T} r(t)p(t+\epsilon T)^* dt \tag{17}$$

$$\lambda(r(t)|H) = \mathbb{E}_{\epsilon}\left\{\mathbb{E}_{\theta_0}\left\{\mathbb{E}_{b_0 b_1 \ldots b_K}\left\{e^{-\frac{(r(t)-s(t))^2}{N_0}}\right\}\right\}\right\} \tag{18}$$

3.2.6 Development of MPSK Classifiers

The development of optimal decision rules tries to simplify the equations usually by changing the order of the expectations, reducing the equations using Taylor approximation and using some other desirable means of simplification such as taking the log-likelihood.

3.2.6.1 Averaging Over the Symbols

For a synchronous MPSK detector, the likelihood function is given by (19). A new parameter r_k is defined as the correlator output (20). The parameter γ is the signal-to-noise ratio at the output of the correlator. The details of this development are shown in Appendix B.

$$\lambda(r(t)|H) = \prod_{k=0}^{K-1} \mathbb{E}_{b_k}\left\{e^{\sqrt{\gamma}\, Re\{\, r_k(\epsilon)\, b_k{}^{*} e^{-j\theta_o}\} - \gamma N}\right\} \tag{19}$$

$$r_k(\epsilon) = \frac{2}{\sqrt{N_0 T_c}} \int_{(k-\epsilon)T}^{(k+1-\epsilon)T} r(t)\; p(t-(k-\epsilon)T)\;\, dt \tag{20}$$

Averaging over an arbitrary symbol provides the standard form of the likelihood function.

$$\lambda(r(t)|H = M) = e^{-NK\gamma} \prod_{k=0}^{K-1} \frac{2}{M} \sum_{m=0}^{M/2-1} \cosh\left(\sqrt{\gamma}\, Re\{\, r_k(\epsilon)\, e^{-\frac{j2\pi m}{M}} e^{-j\theta_o}\}\right) \tag{21}$$

Equation (21) can be conditioned on phase and time offset parameters by taking the expectation over those parameters as shown in (18).

Huang and Polydoros (Huang & Polydoros, 1995) noted in a binary hypothesis test of $\lambda(r(t)|M)$ and $\lambda(r(t)|M')$, where $M \geq M'$ and both M and M' are powers of two, it was possible to reduce (20) into two expressions. The first expression is a common term shared by $\lambda(r(t)|M)$ and $\lambda(r(t)|M')$ as a function of M' and the second term is a function of M. The term provided in (22) will serve as the decision rule for MPSK signals.

$$\begin{aligned} q_M &= ln\{\lambda(r(t)|M)\} - ln\{\lambda(r(t)|M)\} \\ &= ln\left\{\mathbb{E}_{\epsilon}\left\{\mathbb{E}_{\theta_0}\left\{exp\left(\frac{2}{M}\left(\frac{\gamma}{2}\right)^{M/2} Re\left\{\left(\sum_{k=0}^{K-1} (r_k(\epsilon))^M\right) e^{jM\theta_o}\right\}\right)\right\}\right\}\right\} \end{aligned} \tag{22}$$

The term $(r_k(\epsilon))^M$ is the heuristic power law detector for MPSK signals. Given an ideal MPSK signal $r(t) = \; e^{j2\pi\alpha(k)/M}$, the power law detector produces a constant response $r(t)^M = \; e^{j2\pi\alpha(k)}$ in the baseband signal and a fixed tone in the cosine modulated signal.

The rule for the non-coherent detection case is shown in (22). Later on, during the discussion of the development of the spread spectrum detector, we will show that the detection rule is identical. The conclusion is: it is not possible to distinguish a spread spectrum signal from an MPSK signal based on simple assumptions of independent waveform coefficients.

3.2.6.2 BPSK Spread Spectrum Model

The case of spread spectrum signals shows some similarities to the MPSK classifier. In this scenario, we deal with complex waveforms instead of simple pulses. The waveforms belong to a set of mutually orthogonal or quasi-orthogonal waveforms according to (23) and (24). Each waveform will serve as a separable coding channel. The spread spectrum signal is the superposition of modulated waveforms according to (25).

$$p_u(t) = \frac{1}{\sqrt{LT}} \sum_{i=0}^{L-1} c_{ui}\, \psi(t-(i-1)\,T) \qquad 0 \le t \le TL \tag{23}$$

$$\int_0^T p_u(t) p_v(t) dt = L\delta_{uv} \tag{24}$$

$$s(t) = \sqrt{\frac{E}{LT}} e^{j\theta_o} \sum_{u=0}^{U-1} \sum_{k=0}^{K-1} b_{uk} \sum_{i=0}^{L-1} c_{ui}\, \psi(t-kLT-(i-1)\,T) \tag{25}$$

Under MPSK hypothesis with unknown phase and time offsets, the likelihood takes the following form:

$$\lambda(r(t)|H) = E_\epsilon \left\{ E_{\theta_c} \left\{ E_{b_{00} b_{01} \dots b_{U-1,K-1}} \left\{ E_{c_{00} c_{01} \dots c_{U-1,L-1}} \left\{ e^{\frac{2}{N_0} \int_0^{NT_S} Re\{ r(t) \cdot s^*(t-\epsilon) \} dt - \frac{1}{N_0} \int_0^{NT_S} |s(t-\epsilon)|^2 dt} \right\} \right\} \right\} \right\} \tag{26}$$

The energy term of the received signals was removed from the equation because it is constant. The transmit signal term is kept because the amplitude of the spread spectrum signal is no longer constant. It is easy to note the complexity of this expression. The expectation must be taken over each possible combination of waveform coefficients and symbols. In the case of binary spread spectrum signals, the integral would require $2^{KL} + 2$ integrations.

We started a simple development by averaging over simplistic assumptions. We begin assuming that the waveform coefficients and symbols are statistically independent. This would allow having an equation similar to (19), where the expectation of product equals the product of expectations.

3.3 Development of LRT for BPSK-SS

3.3.1 Simple Averaging Over the Waveform Coefficients and Symbols

A synchronous BPSK spread spectrum detector, the likelihood function is given by (27). A new parameter r_k is defined as the correlator output (28). The parameter γ is the signal-to-noise ratio at the output of the correlator. The details of this development are shown in Appendix B.

$$\lambda(r(t)|H) \triangleq E_{c_{00}c_{01}\dots c_{U-1,L-1}}\left\{e^{\frac{2}{N_0}\int_0^{NT_S} Re\{\,r(t)\cdot s^*(t-\epsilon)\}dt-\frac{1}{N_0}\int_0^{NT_S}|s(t-\epsilon)|^2 dt}\right\} \tag{27}$$

$$r_{ki}(\epsilon) = \frac{2}{\sqrt{N_0 L T_c}}\int_{(k-1+\epsilon)T}^{(k+\epsilon)T} r(t)\,\psi(t-(k-\epsilon)TL-(i-1)\,T)\;dt \tag{28}$$

Assumption #1: The energy term of the transmitted signal will be treated as a global constant. The distribution of the overall symbol amplitude (29), which is the sum of the symbols times the waveform coefficients, is a binomial distribution.

$$s[m] = \sum_{u=0}^{U-1}\sum_{k=0}^{K-1} b_{uk}\sum_{i=0}^{L-1} c_{ui}\,\delta(m-kL-(i-1)\,) \tag{29}$$

The term is $E_{c_{00}c_{01}\dots c_{U-1,L-1}}\left\{e^{-\frac{1}{N_0}\int_0^{NT_S}|s(t-\epsilon)|^2 dt}\right\}$ is a function of the variance of $s[m]$ which depends on the number of users U and length of the spreading code L. So the only term of interest is the correlation term.

Assumption #2: The waveforms coefficients are statistically independent. The likelihood is simplified to (30).

$$\lambda(r(t)|H) \triangleq \prod_{u=0}^{U-1}\prod_{i=0}^{L-1} E_{c_{ui}}\left\{e^{\frac{2}{N_0}\int_0^{NT_S} Re\{r(t)\cdot s^*(t-\epsilon)\}dt}\right\} \tag{30}$$

It was found that the assumption is not useful. Averaging over statistically independent symbols will also yield a product of terms as in (31).

$$\lambda(r(t)|H) \triangleq \prod_{u'=0}^{U-1}\prod_{k=0}^{K-1} E_{b_{u'k}}\left\{\prod_{u=0}^{U-1}\prod_{i=0}^{L-1} E_{c_{ui}}\left\{e^{\frac{2}{N_0}\int_0^{NT_S} Re\{r(t)\cdot s^*(t-\epsilon)\}dt}\right\}\right\} \tag{31}$$

A lengthy development reveals that using a second assumption results is equivalent to the BPSK likelihood. The conclusion is: the assumption of statistically independent coefficients does not allow differentiating between a BPSK signal and a BPSK spread spectrum signal.

3.3.2 Weighted Average over the Waveform Coefficients and Symbols

The proposed approach uses a statistically dependent distribution for averaging over the waveform coefficients. The approach requires us to design a waveform density based on some metric that characterizes spreading sequences. We consider the metric: Total Square Correlation (TSC) defined as the Frobenius norm of the correlation matrix.

Let $C = \{c_{ui}\}$ be the matrix of waveform coefficients row u and column i. Also consider a row vector representation $\vec{c}_u = [c_{u0}, c_{u1}, \dots, c_{u,L-1}]$. The correlation matrix is defined as (32). (Cotae, 2003)

$$TSC(C) = \sum_{u=0}^{U-1}\sum_{u'=0}^{U-1} |\vec{c}_u \vec{c}_{u'}^{H}|^2 \geq U, \quad for\ L \geq U \tag{32}$$

A modification of this metric was used for our density as shown in (33).

$$D(C) = \sum_{u=0}^{U-1}\sum_{\substack{u'=0 \\ u\neq u'}}^{U-1} |\vec{c}_u \vec{c}_{u'}^{H}|^2 \geq 0 \tag{33}$$

For a binary spread spectrum, the coefficients take two real values: $c_{ui} = \pm 1$, so any power also stays in the same range: $c_{ui}^n = \pm 1$. The density for statistically dependent coefficients is (34).

$$e^{-\delta \cdot D(C)} \tag{34}$$

$$D(C) = (U^2 - U)L + 4 \sum_{i=0}^{U-1} \sum_{\substack{j=0 \\ j \neq i}}^{U-1} \sum_{k=0}^{L-1} \sum_{\substack{l=0 \\ k \neq l}}^{L-1} c_{ik} c_{il} c_{jk} c_{jl} \tag{35}$$

The expression of the likelihood (36) is complicated and further analysis will be provided in a future report. In the present discussion, we are going to develop the likelihood function for the simplest case $U = L = 2$.

$$\lambda(r(t)|H)$$
$$= \sum_{c_{00}} \sum_{c_{01}} \dots \sum_{c_{U-1,L-1}} e^{\sqrt{\gamma} \sum_{u=0}^{U-1} \sum_{k=0}^{K-1} \sum_{i=0}^{L-1} Re\{ r_{ui}(\epsilon) b_{uk}^* c_{ui}^* \, e^{-j\theta_o} \} - \gamma \sum_{u=0}^{U-1} \sum_{i=0}^{L-1} b_{uk}^* c_{ui}^* b_{uk} c_{ui} - D(\vec{c}_u)} \tag{36}$$

3.3.3 Simple Case: 2x2 Hadamard Coefficients

The likelihood for detecting 2x2 Hadamard coded spread spectrum was computed in (37) assuming statistical dependency of the waveform coefficients. The results were encouraging as we can see from Figures Figure 1-Figure 3. A derivation of the likelihood was done in *Mathematica*. (See Appendix C)

$$\lambda\big(r(t)|H\colon\{\,c_{ij}=\pm1, U=2, L=2\}\big)$$
$$=\frac{1}{4\cdot 8\cdot(1+e^{2\delta})}\Big(e^{-2\sqrt{2\gamma}}+e^{-2\delta}\cosh\left(2\sqrt{\gamma}\,Re(r_{k0})\right)$$
$$+\,e^{-2\delta}\cosh\left(2\sqrt{\gamma}\,Re(r_{k1})\right)$$
$$+\,e^{2\delta-4\gamma}\left(\cosh\left(2\sqrt{\gamma}\,Re(r_{k0})\right)\cosh\left(2\sqrt{\gamma}\,Re(r_{k1})\right)\right)\Big) \tag{37}$$

$$\delta=\sqrt{2\gamma}$$

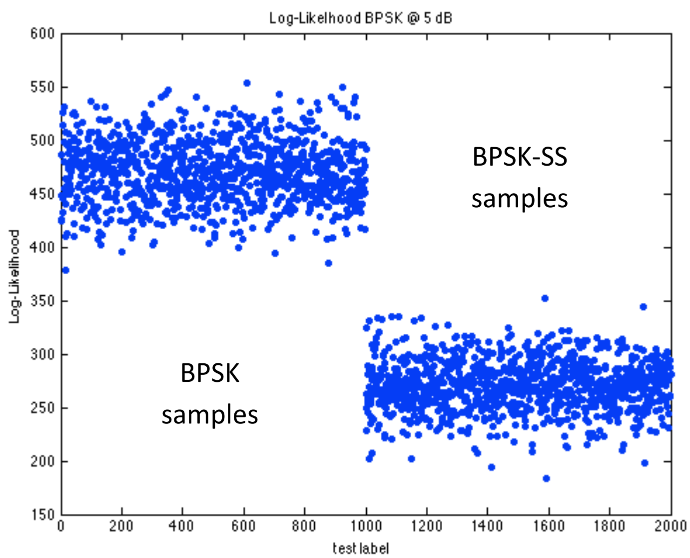

Figure 1. Likelihood of BPSK

The likelihood of the spread spectrum signal can be decomposed in two terms: the individual BPSK terms for r_{k0} and r_{k1}, and the product of both terms. The LRT test gives acceptable results for SNR ranges between 0 and 5 dB. For higher ranges, the test has problems due to overflow (1000 chips used) and the test breaks for levels below -5dB.

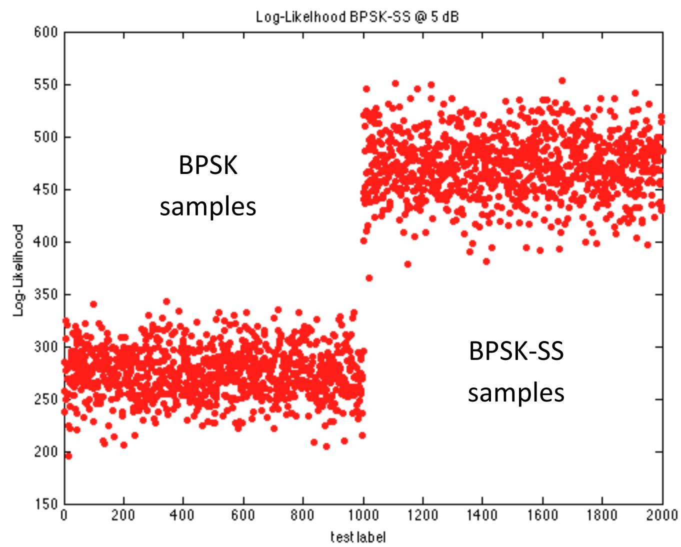

Figure 2. Likelihood of BPSK Spread Spectrum

Figure 1 and Figure 2 show the likelihood for BPSK and BPSK Spread Spectrum. Both likelihoods can be approximated using Gaussian distributions. The empirical receiver operating characteristic (ROC) curve is shown in Figure 3.

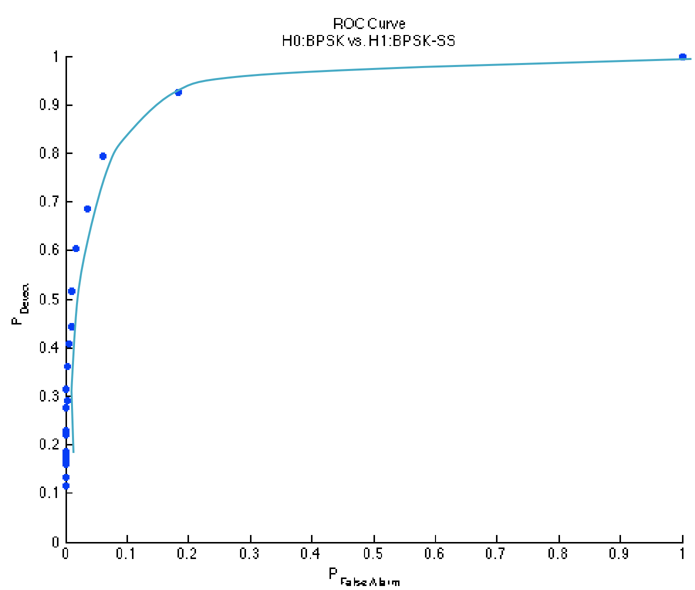

Figure 3. ROC Curve for LRT test
H0: BPSK, H1: BPSK-Spread Spectrum

3.4 Inference Methods for Classification

In order to overcome the complexity of our blind demodulation problem, this paper proposes to modify a stochastic approach for signal demodulation by incorporating the concept of information divergence or "distance" between probability density functions. One of the densities provides a parametric model; the second density describes the ideal modulation type. (See equations (38, (56) and (58)) The models can describe noise or any arbitrary modulation type. An optimization algorithm (54) will adjust the parameters of the distribution to fit various models by minimizing the divergence. The decision is made by selecting the smallest "distance" instead of the likelihood ratio.

The Expectation Maximization approach proposed by Yao (Yao & Poor, 2000) provides an iterative method for estimating the parameters in equation (38). The scheme uses BPSK-SS model with a known number of users. The signal model is:

$$\vec{y} = \sqrt{E} \cdot S \cdot \vec{b} + \vec{n} \tag{38}$$

Where n is the noise, $\vec{b} \in \{\pm 1\}^K$ is the transmitted symbol vector of length K. The matrix S is a $L \times K$ matrix with K users and L the signature size. The S matrix contains the normalized spreading sequences $S = [\,\vec{s}_0, \vec{s}_1, \dots \vec{s}_K\,]$ E is the energy of the signal. The output $\vec{y}$ is the received vector. The process can be described with a correlation matrix R:

$$\mathbb{E}\{\vec{y}\vec{y}^H\} = R = U^H \Lambda U \tag{39}$$

Eigenvalue decomposition provides a means to separate the signal space from the noise space. The highest eigenvalues correspond to the energy of the signals and the associated eigenvectors correspond to the signature vectors that form matrix S.

$$\begin{gathered}\Lambda = \Lambda_\mathrm{s} + \Lambda_\mathrm{n} = \mathrm{diag}(\lambda_0, \lambda_1, \dots, \lambda_\mathrm{K}, \lambda_{\mathrm{K}+1}, \dots, \lambda_\mathrm{L}) \\ \Lambda_\mathrm{s} = \mathrm{diag}(\lambda_0, \lambda_1, \dots, \lambda_\mathrm{K}) \quad \textit{signal space} \\ \Lambda_\mathrm{n} = \mathrm{diag}(\lambda_{\mathrm{K}+1}, \dots, \lambda_\mathrm{L}) \quad \textit{noise space}\end{gathered} \tag{40}$$

The associated eigenvectors U_s to the signal space can be used to estimate the matrix S by means of (41).

$$U_s^H \vec{y} = \ x = \sqrt{E} \cdot S \cdot \vec{b} + \ \sigma \cdot n \tag{41}$$

The conditional likelihood function becomes (42):

$$p(x|\sqrt{E}S) = \ \sum_{k=1}^{2^K} \frac{1}{2^K} \frac{1}{(2\pi\sigma_s)^{K/2}} \ \exp\left(-\frac{1}{2\sigma_s^2} \left\|\vec{x} - \sqrt{E}S\vec{b}_k\right\|^2\right) \tag{42}$$

If $\vec{b}_k$ is a binary vector, then there are 2^K possible combinations of symbols. The EM process adds a new variable to the conditional likelihood. This parameter is known as the missing data i_n. The missing data is a label that identifies the population or cluster where the known variables belong to. In this case, the cluster is any binary pattern formed from numbers 0 to $2^K - 1$.

$$p(x, i_n|\sqrt{E}S) = \ \sum_{n=1}^{N} \frac{1}{(2\pi\sigma_s)^{\frac{K}{2}}} \exp\left(-\frac{1}{2\sigma_s^2} \left\|\vec{x} - \sqrt{E}S\vec{b}_{i_n}\right\|^2\right) p(i_n) \tag{43}$$

The EM method maximizes the expectation of the log-likelihood given the data. This is:

$$\hat{S} = \arg\max_{S} \mathbb{E}_{i_n|y} \left\{\log\left(p(x, i_n|\sqrt{E}S)\right)\right\} \tag{44}$$

A detailed explanation of this algorithm explained in (Yao & Poor, 2000). By using an EM approach, the process of maximizing the expectation eventually that converges to the waveform coefficients in matrix S without dealing with the complexity of the ALRT. According to the paper, the estimation of the waveform coefficient offers a slightly worse performance compared to a minimum square error estimator during the demodulation. The worse performance is expected because MMSE assumes known waveform parameters while in the EM approach the estimation is totally blind.

The development of the EM estimator allows for applying constraints, but these are not explored in the referenced paper. For instance, other possible constraint is to minimize the entropy of the likelihood function (45) for all possible signals. This constraint would force that the results reduce the randomness.

$$\sum_{n=1}^{N} p(x|i_n)\log(p(x|i_n)) = 0 \tag{45}$$

3.4.1 Information Divergence Modulation Classifier

Under the FY11-12 AFOSR Mini-Grant, the concept of information divergence was studied as a symbol estimator of unknown signals. The concept was used as a blind estimator as reference (Ding & Kay, 2011) suggested.

Instead of using a ratio of likelihood functions, the proposed approach uses the Kullback-Leibler (KL) divergence. KL divergence is a measure of "distance" or dissimilarity between two distributions. Given two distributions $\mathrm{u(x,y)}$ and $\mathrm{v(x,y)}$, the Kullback-Leibler divergence is defined by (46):

$$D_{KL}(u||v) = -\int_x \int_y u(x,y)ln\left(\frac{v(x,y)}{u(x,y)}\right)dxdy \tag{46}$$

If two distributions are the same, then it is easy to observe that:

$$D_{KL}(v||v) = 0 \tag{47}$$

The divergence between two distributions is always positive. (Cover & Thomas, 2006)

$$D_{KL}(u||v) \geq 0 \tag{48}$$

The divergence is related to the likelihood ratio. Consider two distributions $\mathrm{v(x,y)}$ and $\mathrm{u(x,y)}$, then the divergence between v and u is:

$$D_{KL}(u||v) = \mathbb{E}_{u(r,H_1)}\{log(\Lambda(r))\} + \mathbb{E}_{u(r,H_1)}\left\{log\left(\frac{p(H_1)}{p(H_0)}\right)\right\} \tag{49}$$

where $\Lambda(\mathrm{r})$ is the likelihood ratio:

$$\Lambda(r) = \frac{u(r|H_1)}{v(r|H_0)} \tag{50}$$

and $\mathrm{E}_{\mathrm{u(r,H_1)}}$ is the expectation operator. Using divergence is an extension of the concepts found in Decision Theory. The divergence includes an average of the logarithm of the likelihood ratio.

KL divergence has been used in the speaker recognition area. In this problem, a feature space vector $\vec{x}$ contains information that is used to discriminate between speakers. Suppose that one has a stochastic model $p_0(\vec{x})$ such as a multivariate Gaussian distribution $\mathcal{N}(\vec{\mu}_0, \Sigma_0)$ using the features of the population as random variables. Suppose that a second Gaussian model describes a target speaker. A way for designing a decision rule $p_1(\vec{x}) = \mathcal{N}(\vec{\mu}_1, \Sigma_1)$ between the two classes consists of using a subset of $\vec{x}$ that maximize the divergence $D_{KL}(p_1||p_2)$ between the two classes. The distance between the data and the model is computed using only the selected features and the class that has the "shortest distance" wins. The concept is similar to the Mahalanobis distance $d_M^2 = (\vec{x} - \vec{\mu}_i)^T \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i)$ used to discriminate between classes. In fact, the KL divergence is equivalent to the Mahalanobis distance when $\vec{\mu}_0 = \vec{\mu}_1$ for unimodal Gaussian distributions (Campbell, 1997).

The proposed modulation classifier uses similar concepts of minimum distance for deciding between classes. The feature space is the complex plane that contains the values of the estimated symbols, each one with an in-phase and quadrature component. The stochastic model is assumed to be Gaussian. The true symbols form a constellation which can be rotated due to phase incoherency between the transmitter and the receiver.

We define our communication channel using the diagram of Figure 4. (Cover & Thomas, 2006) The symbols can be drawn from any arbitrary distribution. It can be circular or rectangular QAM or it can have arbitrary constellations.

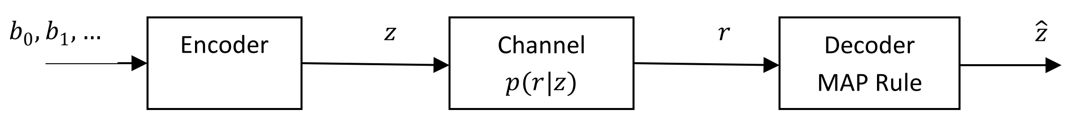

Figure 3. Communication Channel

The probability of a symbol $p(z)$ is assumed to be an unknown variable for computational purposes. This prior knowledge must match the implicit probability of symbols in the template, as it will be explained later.

Table 3. Definitions for EM Approach

Concept:	Mathematical Expression:
Symbol	$z \in S_z$ with S_z being a finite discrete set of complex numbers
Symbol Probability	$p(z)$
Observation	$r \in S_r$ with S_r being a finite set of complex numbers
Noise Power Density	$\sigma^2 = \frac{N_0}{2}$
Likelihood of the observation given a symbol (channel noise)	$p(r\|z) \sim \mathcal{N}(z, \sigma^2)$
Probability of a symbol given the observation	$p(z\|r) = \frac{p(r\|z)p(z)}{p(r)}$
Matching density function	$q(r)$
Sampled Based Kulback-Leibler Divergence	$D_{KL}(u\|\|v) = -\sum_{r \in S_r} \sum_{z \in S_z} u(r,z) ln\left(\frac{v(r,z)}{u(r,z)}\right)$
MAP (Maximum Posterior Rule) for choosing between classes C1 and C2	$p(z\|r, C_1) \underset{C_0}{\overset{C_1}{\gtrless}} p(z\|r, C_2)$

The channel introduces noise of each symbol. The noise model is expressed as a likelihood function. The noise is assumed to be AWGN but this model can be modified to other types of noise if needed.

The correlator in the receiver produces the observation r. For each observed data point there is a probability of a symbol given the observation p(z|r). The maximum posterior probability criterion can be used as a decision rule to compute the detected symbol ẑ.

Bayes' theorem states that:

$$p(z|r)p(r) = p(r|z)p(z) \tag{51}$$

Note that the divergence between $p(z|r)p(r)$ and $p(r|z)p(z)$ is zero:

$$D_{KL}\big(p(z|r)p(r)||\, p(r|z)p(z)\big) = 0 \tag{52}$$

Suppose now that we do not want to use $p(z)$. Instead, we would like to use a template distribution q(r) . This template satisfies the inequality:

$$D_{KL}\big(p(z|r)q(r)||\, p(r|z)p(z)\big) \geq 0 \tag{53}$$

The equality holds only and only if $q(r) = p\big(r;\vec{\theta}\big)$ for a given parameter vector $\vec{\theta}$. If the equality does not hold, there is at least a minimum divergence that can be achieved between the two distributions according to Kullback's Minimum Discrimination of Information Principle (MDI). (Gray, 2001) The parameter vector can include noise, phase and signal energy parameters. In the simulations, only phase incoherency was considered.

The optimal parameters can be determined using the proposed optimization criterion:

$$\vec{\theta}^{opt} = \arg\max_{\vec{\theta}} D_{KL}\left(q(r,z;\vec{\theta})||\, p\big(r,z;\vec{\theta}\big)\right) \tag{54}$$

where $q\big(r,z;\vec{\theta}\big) = \; p\big(z\big|r;\vec{\theta}\big)q(r)$ and $p\big(r,z;\vec{\theta}\big) = \; p\big(r\big|z;\vec{\theta}\big)p(z;\vec{\theta})$. The algorithm can be regarded as a stochastic matched filter.

Once the optimal parameters have been determined for classes with templates $q_1(r)$ and $q_2(r)$, the decision is made either using the conventional maximum posterior probability or using the KL divergence.

3.4.2 Simulations

MPSK models and QAM models were considered for the estimation of the parameters.

3.4.2.1 MPSK Case:

If symbols z belong to the MPSK class, then:

$$z \in \left\{ e^{\frac{j2\pi m}{M}}, m = 0, 1, \dots, M-1 \right\} \tag{55}$$

The likelihood model that describes the channel is the simple Gaussian distribution shown in (56):

$$p\left(r \middle| z;\ \vec{\theta}\right) = \frac{1}{(\pi N_0)^{1/2}} exp\left(-\frac{\left|r - ze^{j\theta}\right|^2}{N_0}\right) \tag{56}$$

The model describes a noisy output with rotated symbols $z\,e^{j\theta}$. The parameter vector consists of two quantities, the phase mismatch and the probability of the symbol:

$$\vec{\theta} = [\theta, \quad p(z)] \tag{57}$$

The template of an arbitrary MPSK hypothesis is:

$$q(r|H_M) = \sum_{l=0}^{M-1} \frac{1}{M(\pi N_0)^{1/2}} exp\left(-\frac{\left|r - e^{\frac{j2\pi l}{M}}\right|^2}{N_0}\right) \tag{58}$$

Note that the template incorporates the true prior distribution $\mathrm{p(z)} = 1/\mathrm{M}$. This is not a necessary requirement, but provided a convenient quick implementation.

Figure 4 to Figure 8 shows the convergence of the algorithm (49) for MPSK signals. The observed data r is shown as green dots. The red dots track the convergence of the data to the model, i.e., $ze^{j\theta} \rightarrow e^{\frac{j2\pi l}{M}}$. The blue dots represent the true symbols and the yellow dots (barely visible) represent the final convergence. The black lines represent the decision boundaries for MPSK types. (See Figure 5 and Figure 6)

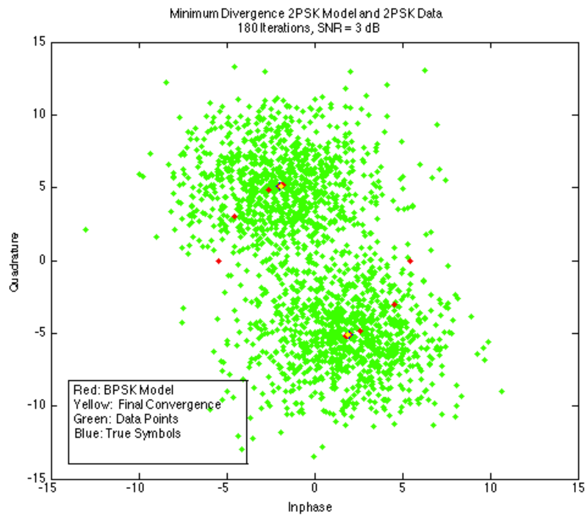

Figure 4. Convergence of the MDI Estimator/Classifier using BPSK Data

In Figure 5, we achieved convergence in 4 iterations. BPSK is the simplest form of MPSK types. Quick convergence is expected even in low SNR levels.

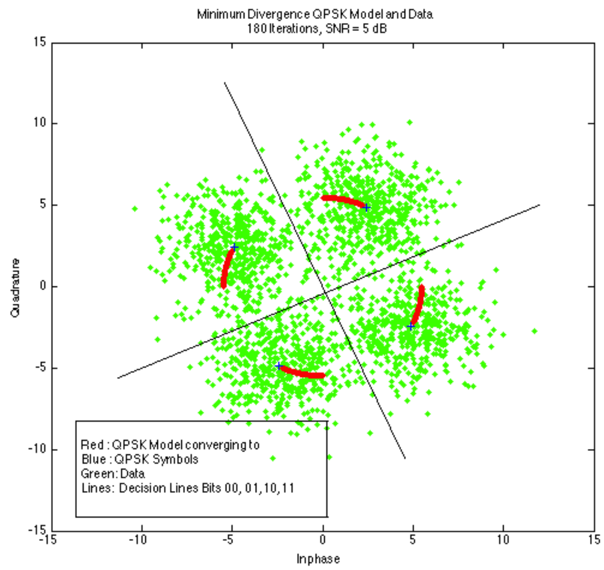

Figure 5. Convergence of the MDI Estimator/Classifier using QPSK Data

In Figure 5, convergence of QPSK data to QPSK model is achieved in approximately 50 iterations. The computational time is still insignificant due to the implementation of the algorithm which is based on simple linear algebra and element by element operations.

In Figure 6, convergence of QPSK data to BPSK model is achieved in approximately 50 iterations. The minimum distance between a QPSK and a BPSK densities is achieved when we distribute the data equally along the decision line. The resulting divergence of this case is greater than the divergence. In such case, the classifier would decide for QPSK instead of BPSK.

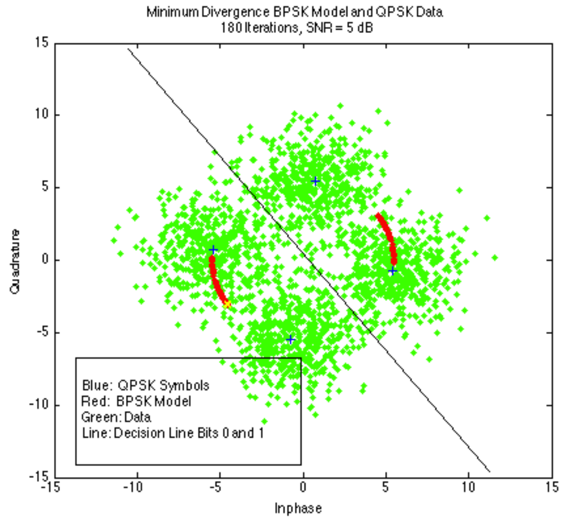

Figure 6. Convergence of the MDI Estimator/Classifier using QPSK Data and BPSK model

The reader might wonder if an algorithm such as Gaussian Mixture Models (GMM) clustering algorithm can perform a similar task. The answer is negative. Differently from the GMM algorithm, the MDI classifier keeps fixed proportions between the symbols (or clusters) and prohibits them from moving independently.

Figure 7 shows a case of a noisy 16PSK signal. The convergence of cases like this one may take several hundred of iterations but running in fractions of seconds. It is very likely that optimal classifiers fail in a case like this one. The only solution to a problem like this is to observe more data.

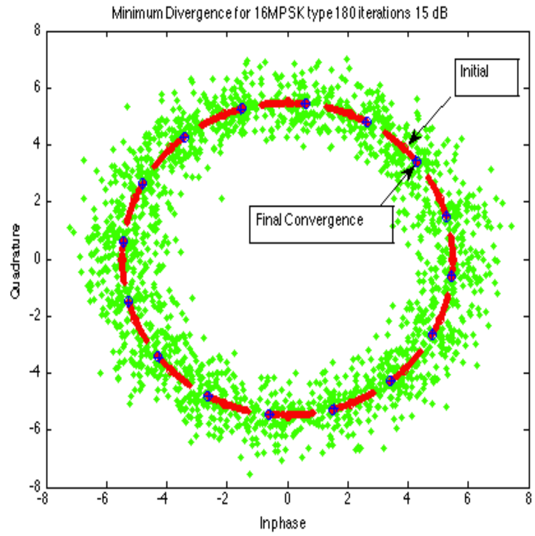

Figure 7. Convergence of the MDI Estimator/Classifier using 16PSK Data

3.4.3 QAM Case

The proposed classifier can be flexible enough to incorporate other constellations such as QAM models. (Figure 8) The only requirement is to provide a template as an input. Other constellations (templates) are admissible by just providing a different template model q(r|H).

The output of the algorithm is the optimal parameter vector and the divergence. With the parameter vector, the likelihoods and priors for each observation can be reconstructed. This is good feature because these variables not only allow deciding the modulation type, buy they provide sufficient information for recovering the symbols and the bitstream.

In the case of complex constellations (QAM-32, QAM-64), the convergence is less reliable due to the multiple local minima that exist. The method is more likely to work on simple constellations like BPSK, QPSK signals and BPSK-SS waveforms.

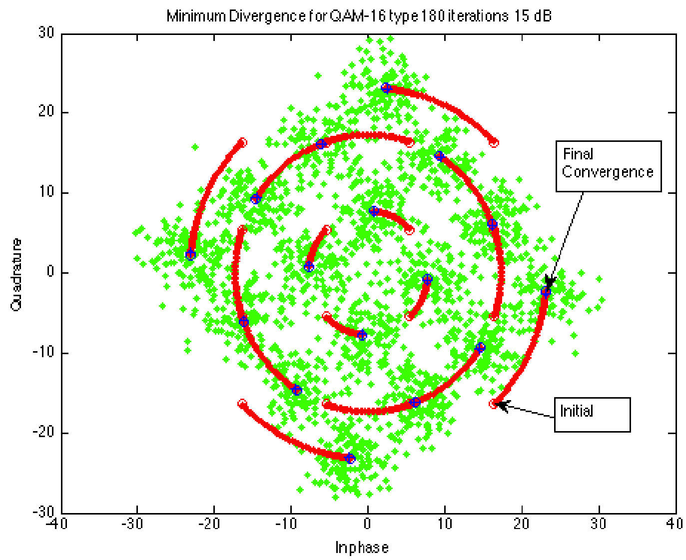

Figure 8. Convergence of the MDI Estimator/Classifier using 16QAM Data

3.4.3.1 Divergence for Waveform Estimation

The benefit of using information divergence is that the method preserves the signal constellation proportions and only allows rotation or scaling. This is a feature the EM method does not have.

4 MIMO Code Design

The purpose of this research study is to investigate the MIMO Space-Time Block Codes (STBC) and provide if possible an improved design. In this report, the available codes will be presented. The best-known code will be identified and study for the purpose of developing better codes.

4.1 Introduction

MIMO is a multiple input multiple output communication systems that transmit information from a multiple antenna transmitter to a multiple antenna receiver in the presence of fading. One of the benefits that make MIMO system attractive is the increasing channel capacity due to the antenna diversity. MIMO systems are currently used in mobile networks such as cell phones and routers, where voice and data requires higher transmissions rates and better quality of service.

The design of In MIMO codes is one important topic of MIMO systems. The codes found in literature can be classified in the groups as follows:

- Layered Spaced Time Coding
- Space Time Block Coding
- Trellis Space Time Coding.

Each code family differs in the degree of complexity versus accuracy. Layered Space Time Coding is an example of simplistic codes with a poor performance compared to trellis codes with the highest complexity and the best accuracy. Block coding provides a trade-off between performance and complexity.

This report is dedicated exclusively to the study of the STBC. The design of STBC does not provide error correction as the Trellis Codes, only mitigation of multipath. The best-known STBC is known as the Golden Code. It was developed around 2003 and since this date no other code with superior performance has been found.

Desirable characteristics that make a good code can be described in terms of the following properties:

- Best coding gain

- Minimum average power
- Full rate
- Lowest complexity

After a literature search, the research was oriented to study and propose a new scheme that exceeds the performance of the Golden Code. For such purpose, we will present in the discussion the idea behind the development of the Golden Code before proposing schemes and testing the possible advantages or disadvantages of such codes. Then, we will proceed to investigate schemes that can provide the properties of good codes. Finally we will make an assessment of the designs and get to conclusions whether there is a better STBC than the Golden Code.

4.2 Space-Time Model

The problem of finding new codes will be constrained to the case of frequency non-selective channels, also known as the narrowband case. We also consider a two-transmit two-receive antenna system in the entire discussion unless it is otherwise specified. In this scenario, the frequency response of the channel coefficients is flat in the bandwidth of interest. The channel matrix in (59-(60) characterizes the channel fading:

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \tag{59}$$

where h_{ij} is the gain of the path between transmitter antenna *j* and receiver antenna *i*. The relationship between the output and the input signals under fading and additive white Gaussian noise is: $Y = H \cdot C + N$ or,

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \cdot \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} + \begin{bmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{bmatrix} \tag{60}$$

where y_{ij} is the received signal *i* at time *j*; c_{ij} is the transmitted symbol from receive antenna *i* at time *j*; and n_{ij} is the noise seen by receiver antenna *i* at time *j*.

4.2.1 Criteria for Best Performance

4.2.1.1 Coding Gain

In MIMO systems, the role of the code is to transmit information in such way that it provides immunity to fading in AWGN. A metric of performance is the minimization of the pairwise probability of error (PEP). The PEP given a set of transmitted codes C_k and a detected code $\hat{C}_k$ is defined as:

$$P_e = \frac{1}{K}\sum_{\substack{k=0\\k\neq j}}^{K-1} Pr\left(\hat{C}_j | C_k\right) \tag{61}$$

The analysis (Paulraj, 2003) shows that the PEP a function of the rank r, the number of receive antennas M_R, the product of the eigenvalues $\lambda_k$of the difference $\hat{C}_j - C_k$. An upper bound for the PEP is given by:

$$P_e \leq \frac{1}{2}\left(\prod_{k=0}^{r}\lambda_k\right)^{-M_R}\left(\prod_{k=0}^{r}\frac{\rho}{4M_T}\right)^{-r\cdot M_R} \tag{62}$$

Based on PEP analysis the code gain is given by (63):

$$J = min\frac{1}{2\sqrt{M_T}}\left(\prod_{k=0}^{r}\lambda_k\right)^{1/2r} \tag{63}$$

For code design, we consider a fixed physical system were the number of transmit antenna M_T equals the number of receive antennas M_R. We also considered the case when the number of time periods equals the number of transmit antennas such as the one shown in (60).

Optimizing PEP requires maximizing the code gain. This is achieved in two ways: the matrix difference $\hat{C}_j - C_k$ is full rank. This is the rank criterion. The second is the determinant of the matrix is as large as possible. The combination of the two criteria yields:

$$J = min \frac{1}{2\sqrt{M_T}} \left|det\left(\hat{C}_j - C_k\right)\right|^{\frac{1}{2M_T}} \tag{64}$$

4.2.1.2 Minimum Average Power

The design the best space-time code requires finding a set of codes C_k such that the the set maximizes the coding gain and minimizes the average power. We look at the following Frobenius norm or a matrix trace for optimality.

$$\begin{gathered} max \left|det\left(\hat{C}_j - C_k\right)\right| \\ \min Tr(C_k) \end{gathered} \tag{65}$$

For comparison purposes between codes sets A and B with the same code rate, we normalize the power of A and B and find the highest code gain.

4.2.1.3 Code Rate

Code rate is defined as the number of symbols that can be transmitted in a STBC code. For a 2x2 matrix, the maximum number of symbols is 4. If the code matrix can transmit the maximum number of symbols, then the code is known as full rate FR. If in addition the code has full rank, the matrix is referred as full-rate, full diversity code (FRFD). The best performance STBC codes are FRFD.

4.2.1.4 Code Complexity

A tradeoff between performance and complexity is preferable when the demodulation cost of the receiver exceeds the benefit of overcoming the fading effects. Such cost can be measured in terms of system power consumption, size and/or processing time. Under our discussion, code complexity was not a consideration for the design of the best STBC.

4.3 Existing MIMO Codes

The first task in the search for the best performance codes is to consider the existing code design through the literature. As we mentioned early, MIMO codes found in literature can be classified in the groups.

4.3.1 Layered Space Time Coding

One of the simplest coding schemes is called Horizontal Encoding (HE). In this scheme the transmitted symbols are demultiplexed into separate streams or layers through each transmit antenna in single or multiple time slots. Vertical Encoding (VE) is another variant where the same symbol is spread through all antennas as opposed of being spread through multiple time slots. Diagonal Encoding (DE) transmits the symbols using a different antenna at a different time slot. (Sellathurai & Haykin, 2009) Codes like VE or DE are in general easy to implement, but not accurate due to error propagation.

4.3.2 Space Time Block Coding

Similar to Layered Space Time Codes, these blocks repeats symbols and its complex conjugates in a way that the code achieves full diversity.

4.3.2.1 Alamouti Codes

Alamouti proposed a 2x2 coding scheme that meets full diversity and simple decoding. (Blahut, 2003) The code rate is one symbol per time slot. The design proposed design was extended higher dimensionalities. Its code form can be expressed as:

$$C = \begin{bmatrix} x_1 & x_2 \\ x_2^* & x_1^* \end{bmatrix} \tag{66}$$

Alamouti codes are linear codes that are orthogonal according to:

$$C \cdot C^H = (|x_1|^2 + |x_2|^2)\, I \tag{67}$$

This property facilitates the simplification of the maximum likelihood demodulator:

$$\hat{C} = \arg\min_{C} \|Y - H \cdot C\|_2^2 \tag{68}$$

$$\hat{C} = \arg\min_{C} -2\sqrt{\frac{\rho}{M_T}} Tr\{Y \cdot H^H \cdot C\} - \frac{\rho}{2}(|x_1|^2 + |x_2|^2)\|H \cdot H^H\|_2^2 \tag{69}$$

Further simplification allows expressing the demodulator by expressing the demodulator in separable functions of x_1 and x_2.

Mathematically speaking, the Alamouti code is a Hamiltonian Quaternion isomorphism. A Hamiltonian Quaternion is an algebraic structure known as ring that obeys associative and distributive laws under addition and multiplication while preserving the same algebraic structure of the operands. This discovery was found latter, with the development of FRFD codes.

4.3.2.2 Quasi-Orthogonal Codes

Several variants of Alamouti code for higher dimensions have been proposed (Sharna & Papadias, 2003). These code matrices are not orthogonal but the demodulator can be expressed in separable functions. For higher dimensionality the code rate of an Alamouti and quasi-orthogonal matrices decrease, so the transmission becomes less efficient.

4.3.2.3 Algebraic Codes

These codes are developed based on algebraic number theory. The goal of these codes is to ensure full diversity and full rate.

4.3.2.4 Golden Code

Golden Codes were presented by (Belfiore & Viterbo, 2004) as cyclic division algebra, but other isomorphic forms of the code were already known. The Golden Code can transmit 4 symbols in two time slots and ensure full-diversity. So far, it is claimed that the Golden Code has the best performance of all known codes.

$$C = \begin{bmatrix} x_0 & x_1 \\ \gamma\sigma(x_0) & \sigma(x_1) \end{bmatrix} \qquad x_0 = a + \theta b \quad x_1 = c + \theta d \tag{70}$$

4.3.2.5 Perfect Space-Time Codes

Perfect codes are an extension of the 2x2 Golden Codes for higher dimensions.

$$C = \begin{bmatrix} x_0 & x_1 & x_2 & x_3 \\ \gamma\sigma(x_0) & \sigma(x_1) & \sigma(x_2) & \sigma(x_3) \\ \gamma^2\sigma(x_0) & \gamma\sigma(x_1) & \sigma(x_2) & \sigma(x_3) \\ \gamma^3\sigma(x_0) & \gamma^2\sigma(x_1) & \gamma\sigma(x_2) & \sigma(x_3) \end{bmatrix} \tag{71}$$

It has been said that cyclic algebra codes provides full-rate linear codes while cyclic division algebras provide full diversity codes.

4.3.2.6 Trellis Codes

Trellis codes offer the best performance at the expense of a complex implementation. They are also referred as convolutional codes. Due to the complexity of the design, these codes are excluded from this report.

4.4 Group Theory and Golden Codes

The concept of an algebraic structure describes a numeric set such as numbers, vectors, polynomials, matrices of an invariant form that defines one or more operations such like addition, subtraction, multiplication and division.

Groups are algebraic structure under addition and subtraction. The operands belong to the same domain and the result stays in the domain and preserves the same form of its operands. The operands obey associative, distributive and identity laws under addition/subtraction or multiplication. Examples of groups are 2x2 matrices defined in the real or complex domain.

$$\begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} = \begin{bmatrix} a_3 & b_3 \\ c_3 & d_3 \end{bmatrix} \qquad (72)$$

$$a_i, b_i, c_i, d_i \in \mathbb{Z}$$

Rings are algebraic structures under addition, subtraction and multiplication. As in groups, the operands belong to the same domain and the result stays in the domain and preserves the same form of its operands. It rings obey to associative, distributive and identity laws under addition, subtraction and multiplication. Examples of rings are quotient rings of polynomials, modular arithmetic and triangular matrices.

$$\begin{bmatrix} a_1 & b_1 \\ 0 & d_1 \end{bmatrix} \cdot \begin{bmatrix} a_2 & b_2 \\ 0 & d_2 \end{bmatrix} = \begin{bmatrix} a_3 & b_3 \\ 0 & d_3 \end{bmatrix} \qquad (73)$$

$$a_i, b_i, c_i, d_i \in \mathbb{R}$$

Ideals are subsets of rings with special properties. The product of any element of a ring and ideals results in an ideal. Examples of ideals are the set of even numbers. Any real number multiplied by an even number becomes an even number.

$$\begin{gathered} I = 2\mathbb{Z} = \{2 \cdot a | a \in \mathbb{Z}\} \\ a \cdot b = c \\ if\ a \in I, b\ \in \mathbb{Z}, then\ c \in \mathbb{Z} \end{gathered} \tag{74}$$

Fields are algebraic structures under addition, subtraction, multiplication and division. As in all the other algebraic structures, the result of an operation stays in the domain of the operands.

4.4.1 Isomorphism and Homomorphism

Algebraic structures can be mapped into other algebraic structures. We talk about isomorphism when there is a one-to-one transformation between structures. Homomorphism is similar to isomorphism except that several element of an algebraic structure can be mapped into a single element of the second structure.

4.4.2 Why Group Theory?

In STBC design, we look for matrices with full rank and full rate. Full rank implies the existence of a multiplicative inverse. In addition to this, we require that the difference between any two coding matrices is non-singular. These two features can be found in rings.

4.4.2.1 Example of Alamouti Code

Alamouti codes are an isomorphism of Hamiltonian quaternions. The algebraic structure defines a basis of one real and three complex quantities { *1, I, J, K* } that obeys to the following laws:

$$I \cdot J = -J \cdot I, J \cdot K = I, K \cdot I = J\ and\ I \cdot J \cdot K = -1 \tag{75}$$

Any quaternion can be express as linear combination of these four group elements:

$$q = a_1 + a_2 \cdot I + a_3 \cdot J + a_4 \cdot K \tag{76}$$

The sum of quaternions is also a quaternion:

$$
\begin{aligned}
q_1 &= a_1 + a_2 \cdot I + a_3 \cdot J + a_4 \cdot K \\
q_2 &= b_1 + b_2 \cdot I + b_3 \cdot J + b_4 \cdot K \\
q_1 + q_2 &= a_1 + b_1 + (a_2 + b_2) \cdot I + (a_3 + b_3) \cdot J + (a_4 + b_4) \cdot K
\end{aligned}
\tag{77}
$$

The product of a quaternion is a quaternion:

$$
\begin{aligned}
q_1 \cdot q_2 = {} & (a_1 b_1 - a_2 b_2 - a_3 b_3 - a_4 b_4) + (a_1 b_2 + a_2 b_1 + a_3 b_4 - a_4 b_3) \cdot I \\
& + (a_1 b_3 - a_2 b_4 + a_3 b_1 + a_4 b_2) \cdot J \\
& + (a_1 b_4 + a_2 b_3 - a_3 b_2 + a_4 b_4) \cdot K
\end{aligned}
\tag{78}
$$

A quaternion conjugate and inverse is defined as:

$$\bar{q}_1 = a_1 - a_2 \cdot I - a_3 \cdot J - a_4 \cdot K \tag{79}$$

such that

$$\bar{q}_1 \cdot q_1 = a_1^2 + a_2^2 + a_3^2 + a_4^2. \tag{80}$$

An important isomorphism of the quaternion is the 2x2 matrices.

$$q_1 = (a_1 + a_2 \cdot I) + (a_3 + a_4 \cdot I) \cdot J \tag{81}$$

Substituting base *I* by the imaginary number *i* one can map quaternions into matrices:

$$
\begin{aligned}
q_1 \to G_1 &= \begin{bmatrix} a_1 + a_2 i & a_3 + a_4 i \\ -a_3 + a_4 i & a_1 + a_2 i \end{bmatrix} \\
q_2 \to G_2 &= \begin{bmatrix} b_1 + b_2 i & b_3 + b_4 i \\ -b_3 + b_4 i & b_1 + b_2 i \end{bmatrix} = \begin{bmatrix} y_1 & y_2 \\ -y_1^* & y_2^* \end{bmatrix}
\end{aligned}
\tag{82}
$$

There is a mapping between addition and multiplication operations that can be verified.

$$
\begin{gathered}
q_1 + q_2 \rightarrow G_1 + G_2 \\
q_1 \cdot q_2 \rightarrow G_1 \cdot G_2
\end{gathered} \tag{83}
$$

A study of Group Theory revealed that Alamouti matrices are an isomorphic group that obeys the same rules of complex quaternions.

4.4.2.2 Extension of Algebraic Structures

Extending an algebraic structure means that there is a broader structure such as a group, ring or field that includes the elements of the structure. Consider the case of the set of real integers Z. The set forms a ring. This field can be extended to a more general class that includes real and complex integers denoted as $\mathbb{Z}[i]$ also known as Gaussian Numbers.

$$
\mathbb{Z} \rightarrow \mathbb{Z}[i] \tag{84}
$$

An element of the extended structure can be represented using the basis *{ 1, i }*. Any number in the ring can be expressed as a pair of two real coefficients:

$$
\begin{gathered}
x = a + b \cdot i \\
x = a + b \cdot i = (a, b), \qquad a, b \in \mathbb{R}
\end{gathered} \tag{85}
$$

A rational number is defined as a number that can be expressed in terms of two integers: a numerator and a denominator.

$$
Q = \{a/b \mid a, b \in \mathbb{Z}\} \tag{86}
$$

The set of rational numbers forms a field. The field can be extended to the domain of complex integer numbers.

$$
\mathbb{Q} = \{a/b \mid a, b \in \mathbb{Z}[i]\} \tag{87}
$$

The extended field holds four real integers. It can be extended further to include the set of algebraic numbers. Algebraic numbers are defined as any possible root of any arbitrary rational polynomial.

$$\overline{\mathbb{Q}} = \{x \mid P(x) = 0, P(x)\ rational\ polynomial\} \tag{88}$$

A subset of the algebraic numbers is known as the set of quadratic numbers.

$$\overline{\mathbb{Q}}[i] = \{a + \sqrt{d} \cdot b, a, b \in \mathbb{Q}[i]\} \tag{89}$$

A quadratic number can "store" eight integers. Thus by extending the field one can "store" more integer coefficients in a single number structure. In the case of the Golden Code, a quaternion is build using two quadratic numbers. Thus, the Golden Code can "store" 16 integer numbers in a 2x2 matrix structure. This allows getting full rate and full diversity code.

The matrix in (70) forms a division algebra, i.e., a non-commutative division field. The matrix is isomorphic with the cyclic algebra defined as:

$$\begin{aligned} A &= L_1 + e \cdot L_2 \\ a \cdot e &= e \cdot \sigma(a) \end{aligned} \tag{90}$$

The value γ must be a non-relative norm and the function $\sigma(x)$ defines commutation rules for multiplication.

4.4.3 Algebraic Tools for Code Design

In the literature, two main theorems have been use in the design of MIMO codes: the Lindemann Theorem and algebraic number theorem found in (Perlis, 1952), (Paulraj, 2003) and (Ma & Giannakis, 2003).

4.4.3.1 Lindemann Theorem

Given any distinct algebraic numbers $\alpha_1, \ldots \alpha_m$ and non-zero algebraic numbers $a_1, \ldots, a_m$, then the following inequality holds:

$$\sum_{i=1}^{m} a_i \cdot e^{\alpha_i} \neq 0 \tag{91}$$

4.4.3.2 Decomposition of the Extension Field

For a field generator θ in field F, any element β can be uniquely expressed as a polynomial in θ with a_i coefficients in F

$$\sum_{i=1}^{m} a_i \cdot \theta^i = \beta \tag{92}$$

This theorems implies that element β=0, then a_i=0, but $\beta \neq 0$ implies that there is at least one coefficient a_i that is non-zero.

These theorems look similar, but they are mathematically different. The first theorem is based in the distinction between algebraic and transcendental numbers. A transcendental number cannot be expressed as the root of rational polynomials. The second theorem is strictly based in algebraic numbers.

4.4.3.3 Available Tools for Code Construction

In the search for mathematical tools, one should be aware of the existence of isomorphism of Golden Code. Those isomorphisms may include similarity transformations, rotations, scaling, linear mappings and permutations. These forms do not constitute a new code improvement.

4.4.3.4 Matric Polynomials

Matric polynomials of degree m may take the form:

$$\sum_{i=1}^{m} a_i \cdot A^i = A \tag{93}$$

As long the matrix A is full rank and the symbols a_i are not equal to the coefficients of the characteristic equation, the determinant of A will not be zero. A careful selection of a constellation lattice (or equivalently coefficients a_i) may result in matrices with non-vanishing determinants. This construction was not explored in this study.

4.4.3.5 Companion Matrices

Companion Matrices are related to matric polynomials. Code constructions with non-vanishing determinants may be of the form:

$$C(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = log \prod_{j=1}^{4} (A_j - \lambda_j I) \tag{94}$$

The design of C must be such that the difference between any code has non-vanishing determinants. This construction was not explored in this study.

4.4.3.6 Variations of the Golden Code

Another explored form was matrices with determinants similar to the Golden Code. For the Golden Code, the determinant is given by:

$$\begin{gathered} \det(C) = f(x_1, x_2) - i \cdot f(x_3, x_4) \\ f(a, b) \triangleq (x_1^2 - x_1 x_2 - x_2^2) \\ i \cdot f(c, d) \triangleq (x_3^2 - x_3 x_4 - x_4^2) \end{gathered} \tag{95}$$

We note that the determinant has two quadratic multinomials of the same form and one of them is rotated 90 degrees (multipled by the imaginary constant) in the complex plane. Consider the domain of a and b to be the set of Gaussian integers in some rectangular configuration as shown in Figure 9. If we examine the set of points generated by evaluating the multinomial $f(a, b)$ using all admissible values of a and b, we will note that the co-domain are Gaussian integers and there are some empty spaces corresponding to Gaussian integers that never occurs. (The codomain is shown as blue points in Figure 10) We also notice that the multinomial codomain of $i \cdot f(c, d)$ falls exactly in these empty places as shown in Figure 10. The minimum value of the determinant of (82) is determined by the minimum distance between the blue and red points. This observation will be used in the search of new codes.

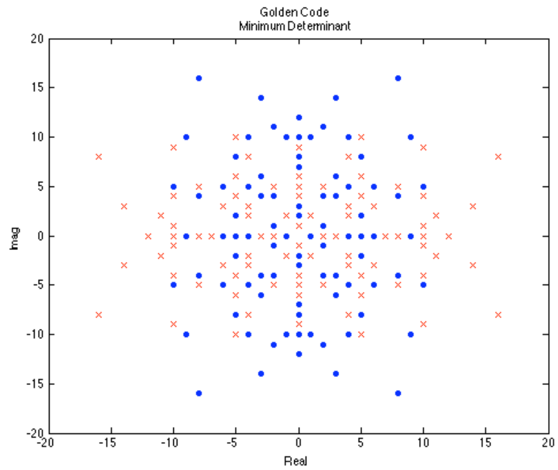

Figure 9. Complex Plane Showing Co-Domain of *f(a,b)* in Blue and *i* f(c,d)* in Red

4.5 Search for New Codes

An interesting question in the code design is: is it possible to find codes similar to Golden Codes that accommodate more than four symbols? The answer turned to be affirmative in this research. Unfortunately, the codes that were found have an average power that is several times higher than the Golden Code.

The code structure of the new proposed codes is similar to (95), but it can accommodate six symbols as shown in:

$$\det(C) = \begin{vmatrix} \alpha_1 s_1 + \alpha_2 s_2 + \alpha_3 s_3 & \alpha_1 s_4 + \alpha_2 s_5 + \alpha_3 s_6 \\ i(\beta_1 s_4 + \beta_2 s_5 + \beta_3 s_6) & \beta_1 s_1 + \beta_2 s_2 + \beta_3 s_3 \end{vmatrix}$$

$$\det(C) = f(s_1, s_2, s_3) + i \cdot f(s_4, s_5, s_6) \tag{96}$$

$$\begin{aligned} f(s_1, s_2, s_3) &= (\alpha_1 s_1 + \alpha_2 s_2 + \alpha_3 s_3)(\beta_1 s_1 + \beta_2 s_2 + \beta_3 s_3) \\ &= a_1 s_1^2 + a_2 s_2^2 + a_3 s_3^2 + a_{12} s_1 s_2 + a_{13} s_1 s_3 + a_{23} s_2 s_3 \end{aligned}$$

It was possible to find such multinomial by doing a random search.

$$f(s_1, s_2, s_3) = s_1^2 + s_2^2 - s_3^2 + (17 + 13j)s_1 s_2 + (13 - 13j)s_1 s_3 + (-13 + 17j)s_2 s_3 \tag{97}$$

$$|Re(s_k)| \leq 3 \; and \; |Im(s_k)| \leq 3$$

A plot of the co-domain of (97) is shown in Figure 10, where none of the rotated points in red overlap with the blue points. The problem is that the multinomial does not factor as a product of linear multinomials. Factoring these terms results in a system of equations that is overdetermined. No solution has been found for this particular case. Without this factorization, the multinomial is useless as a code.

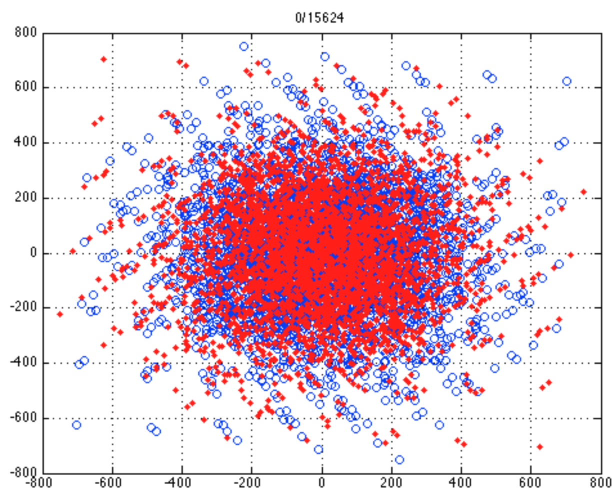

Figure 10. Complex Plane Showing Co-Domain of *f(s1,s2,s3)* in Blue and *i* f(s4,s5,s6)* in Red

The next step would be to find a multinomial form that results in an overdetermined system, but has at least one solution. The form of this code can be expressed as:

$$\alpha_1 = \frac{a_{12} \pm \sqrt{a_{12}^2 - 4a_1 a_2}}{2\, a_2} \alpha 2$$

$$\alpha_3 = \frac{a_{23} \pm \sqrt{a_{23}^2 - 4a_2 a_3}}{2\, a_2} \alpha 2 \tag{98}$$

Equation (98) is a solution obtained by excluding some possible equations of the overdetermined system. Some codes have been found by using conjectures about the form of the solution rather than using a searching algorithm. These codes are shown in Table 4.

Table 4. 2x2 Codes with 6 Symbols

Code	Coefficients	Average Power*
A	$\alpha's: s_1 - \frac{1-4i+\sqrt{5}\sqrt{-2-i}}{2}s_2 + \frac{1+2i+\sqrt{-2+i}}{2}s_3$ $\beta's: s_1 - \frac{1-4i-\sqrt{5}\sqrt{-2-i}}{2}s_2 + \frac{1+2i+\sqrt{-2+i}}{2}s_3$	15767
B	$\alpha's: s_1 + \frac{1-3i+\sqrt{5}\sqrt{2-i}}{2}s_2 + \frac{1+2i+\sqrt{-2+i}}{2}s_3$ $\beta's: -s_1 + \frac{3-4i-\sqrt{5}\sqrt{-2-i}}{2}s_2 + \frac{1+2i+\sqrt{-2+i}}{2}s_3$	14587
C	$\alpha's: s_1 + \frac{6+2\sqrt{5}\sqrt{-2-i}}{2}s_2 + \frac{-1-\sqrt{-2+i}}{2}s_3$ $\beta's: s_1 + \frac{6-2\sqrt{5}\sqrt{-2-i}}{2}s_2 + \frac{-1+\sqrt{-2+i}}{2}s_3$	7488
Golden	$\alpha's: s_1 + \frac{1+\sqrt{5}}{2}s_2$ $\beta's: s_1 + \frac{1-\sqrt{5}}{2}s_2$	200

*After normalizing code gain.

The codes found appear to be related to the golden ratio. The term $\sqrt{5}$ keep reappearing in the codes found. If there is such code that exceeds both performance and rate of the Golden Code, then such code must belong to some sort of unknown ring in the complex domain. The ring, which acts as a place holder for numbers, would be capable of storing more integer coefficients such as (99). One may consider the ring O_K of the set K defined as:

$$K = \mathbb{Q}[i, a^{1/m}, a^{2/m}, \dots, a^{(m-1)/m}] \tag{99}$$

The goal would be to find a scheme that stores more integers while keeping the determinant and power comparable to those ones of the Golden Code. This type of approach will be explored as part of future work.

4.5.1.1 Lindemann Theorem

In this study, we explored the possibility of finding codes of the form:

$$C = \begin{bmatrix} a \cdot e^{\alpha} + b \cdot e^{\beta} & c \cdot e^{\gamma} + d \cdot e^{\delta} \\ i(c \cdot e^{\gamma} + d \cdot e^{\delta}) & a \cdot e^{\alpha} + b \cdot e^{\beta} \end{bmatrix} \tag{100}$$

by conducting an exhaustive search. The determinant of the matrix is a linear combination of transcendental numbers and the Lindemann theorem applies.

A common problem with all these codes is that the determinant approaches to zero when the constellation size grows. Although, Lindemann theorem and field extension decomposition deal with different set of numbers: transcendental and algebraic, it is always possible to find a transcendental number sufficiently close to an algebraic number (Niven, 1953). A number theory theorem states that any irrational number can be approximated with infinitely many rational numbers h/k such that:

$$\left|\xi - \frac{h}{k}\right| < \frac{1}{\sqrt{5}k^2} \tag{101}$$

Finding a good approximation of an irrational number is what causes the determinant to go to zero. If we consider only the diagonal elements of the proposed structure (100), then:

$$\det(C) = a^2 e^{2\alpha} - b^2 e^{2\beta} = 0 \tag{102}$$

Solving for zero yields:

$$\frac{a^2}{b^2} = \frac{e^{2\beta}}{e^{2\alpha}} \tag{103}$$

but under the assumption that a and b are integers, it follows that:

$$\left|\xi - \frac{a^2}{b^2}\right| < \frac{1}{\sqrt{5}k^4}, \qquad \xi = e^{2(\beta-\alpha)} \tag{104}$$

Within the approximations of an irrational number, one must find the worst approximation of rational numbers. For a case of two symbols a and b, it is easy to find values that maximizes the determinant. But for the case of four symbols, it the task is extremely difficult. A *Matlab* simulation shown in Figure 11 shows this fact.

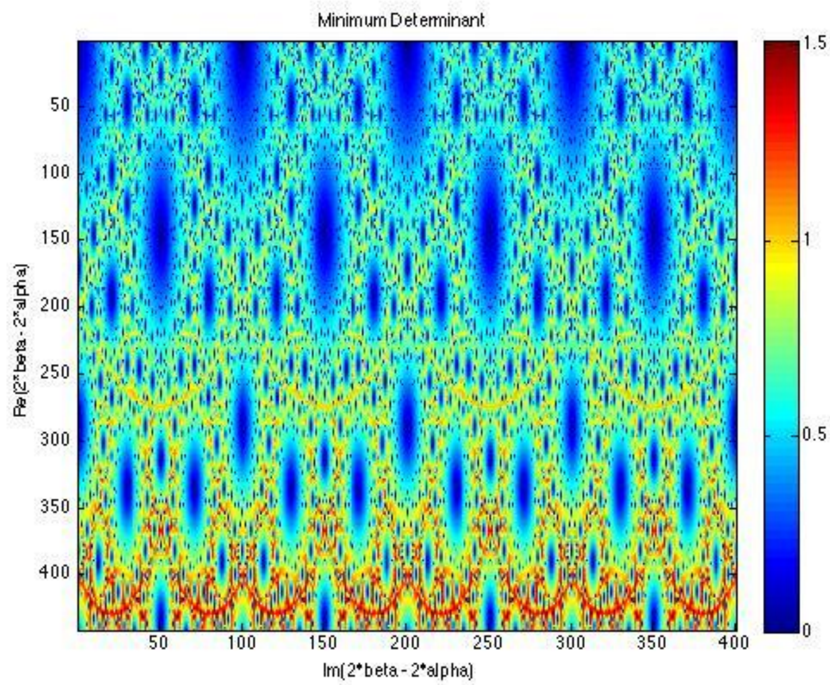

Figure 11. Minimum Determinant of Code based on Lindemann Theorem (102)

In Figure 11, the imaginary and real part of the exponent has been plotted. The real part goes from -0.3 to 0.88 using 500 steps (shown in the figure) and the imaginary part goes from 0 to 2π using 400 steps. The color plot shows the magnitude of the determinant. There is possible to find codes (represented in yellow and red) using Lindemann Theorem such that the determinant is greater or equal than one. Finding full rate codes (using coefficients a, b, c and d) is far more difficult task.

5 Conclusions

Two methods for classification of modulation types were developed: LRT for spread spectrum and EM based for QAM and MPSK.

The first method, which is a classical LRT, provided the desired results when using a modified version of the TSC metric for averaging waveform coefficients using a 2x2 Hadamard matrix. *Mathematica* provided over 1000 exponential terms for this simple case. The LRT test for waveforms using a 4x4 Hadamard matrix is extremely complicated. As a future development, we would attempt to find either a recursive formula based on 2x2 Hadamard codes or deducing some empirical approximation of the LRT.

The second classification method provided a numerical approximation of the symbols on conventional modulation types such as MPSK and QAM. The convergence of the algorithm seems to be an issue for complex QAM constellations. As a future development, we will attempt to improve convergence using Newton's method instead of Gradient Descend. The method will also be extended to BPSK-SS following a Yao's proposed EM algorithm. (Yao & Poor, 2000)

As part of the effort, we also investigated the design of codes that could outperform the current best-known code: the Golden Code. It was possible to find finite codes with higher rates, but also with a higher average power. Future work will include the development of algorithms that conduct a constraint search of these new codes. The goal will be looking for codes with higher spectral rates and similar average power to the Golden Code.

6 References

Belfiore, J., & Viterbo, E. (2004, June). The Golden Code: A 2x2 Full-Rate Space-Time Code with Non-Vanishing Determinants. *IEEE ISIT*, p. 310.

Blahut, R. (2003). *Algebraic codes for Data Transmission.* Cambridge University Press.

Campbell, J. (1997, September). Speaker Recognition Tutorial. *Proceedings of IEEE*, pp. 1437-1462.

Cotae, P. (2003, March). Optimal Sequences and Total Weight Square Correlation of Synchronous CDM Systems. *IEEE Wireless Communication and Networking*, pp. 947-952.

Cover, T., & Thomas, J. (2006). *Elements of Information Theory.* John Willey & Sons Inc.

Ding, Q., & Kay, S. (2011, August). Maximum Likelihood Estimator Under a Misspecified Model with High Signal to Noise Ratio. *IEEE Transactions on Signal Processing*, pp. 4012-4016.

Dobre, O. (2004, Oct). Modulation Classification in Fading Channels using Antenna Arrays. *MILCOM*, pp. 211-217.

Dobre, O., Abdi, A., Bar-Ness, Y., & Su, W. (2007). Survey of Automatic Modulation Classification techniques: Classical Approaches and New Trends. *IET Communications*, pp. 137-156.

Gray, R. (2001). Minimum Discrimination of Information Clustering: Modeling and Quantization with Gaussian Mixtures. *IEEE Image Processing Procedings* (pp. 14-17). IEEE.

Huang, C., & Polydoros, A. (1995, February). Likelihood Methods for MPSK Modulation Classification. *IEEE Transactions on Communications*, pp. 1493-1504.

Kim, K., & Polydoros, A. (1988, October). Digital Modulation Classification: The BPSK vs. QPSK case. *MILCOM*, pp. 431-436.

Lifdtke, F. (1984, August). Computer Simulation of an Automatic Classification Procedure for Digitally Modulated Comm Signals with Unknown Parameters. *Signal Processing*, pp. 331-323.

Ma, X., & Giannakis, G. (2003, November). Full-Diversity Full Rate Complex Field Space-Time Coding. *IEEE Transactions on Signal Processing*, pp. 2917 - 2930 .

Niven, I. (1953). *Irrational Numbers.* Mathematical Association of America.

Panagioutou, P., Anastasoupoulos, A., & Polydoros, A. (2000). Likelihood Ratio Test for Modulation Classification. *MILCOM*, pp. 670-674.

Paulraj, A. (2003). *Introduction to Space time Wireless Communications.* Cambridge Press.

Perlis, S. (1963). *Irrational Numbers.* John Willey and Sons.

Sapiano, P., & Martin, J. (1996, October). Maximum Likelihood PSK Classifier. *Military Communication Conference*, pp. 1010-1014.

Sellathurai, M., & Haykin, S. (2009). *Space-Time Layered Information Process for Wireless Communications.* John Wiley & Sons Inc.

Sharna, N., & Papadias, C. (2003, March). Improved Quasi-Orthogonal Codes Through Constellation Rotation. *IEEE Transactions on Communications*, pp. 332-335.

Soliman, S., & Sue, S. (1992, May). Signal Classification Using Statistical Moments. *IEEE Transactions on Communications*, pp. 908-916.

Yao, Y., & Poor, V. (2000, September). EM-Based Blind Demodulation of Synchronous CDMA. *IEEE 6th Symposium on Spread Spectrum Tech.*, pp. 738-742.

Appendix A

The development of the likelihood function for a process starts with an orthogonal series representation.

$$\mathrm{s(t)} = \sum_{\mathrm{i}=0}^{\infty} \mathrm{b_i}\varphi_{\mathrm{i}}(\mathrm{t}) \tag{1}$$

The signal $n(t)$ is an Additive Gaussian Noise (AWGN) process with mean zero and variance $N_0/2$.

$$\begin{gathered} n(t) = \sum_{i=0}^{\infty} c_i \varphi_i(t) \\ E\{c_i\} = 0 \\ E\{c_i c_j\} = \frac{N_0}{2}\delta_{ij} \end{gathered} \tag{2}$$

A theorem from decision theory tells that the sum of zero mean Gaussian processes is Gaussian, so the coefficients c_i produced by the orthogonal basis decomposition are Gaussian random variables.

$$r(t) = s(t) + n(t) = \sum_{i=0}^{\infty} a_i \varphi_i(t) \tag{3}$$

The variables $\{a_i\}$ have mean $\{b_i\}$ and variance $\left\{\frac{N_0}{2}\delta_{ij}\right\}$. The variables $\{a_i\}$ are also independent.

$$\begin{gathered} E\{a_i\} = 0 \\ E\left\{\left|(a_i - b_j)\right|^2\right\} = \frac{N_0}{2}\delta_{ij} \end{gathered} \tag{4}$$

Processes $s(t)$, $n(t)$ and $r(t)$ can be approximated by selecting a finite set of coefficient. A PDF can be constructed if the coefficients are finite N.

A likelihood ratio test can be constructed as follows:

$$\Lambda(r(t)) = \frac{p_N(r(t)|H_A)}{p_N(r(t)|H_B)} = e^{-\sum_{i=0}^{N}\frac{(a_i-b_i)^2-\left(a_i'-b_i'\right)^2}{N_0}} \tag{5}$$

The expression allow us to arbitrary choose any number of coefficients. If we extend the number to infinity, then we have:

$$\boldsymbol{\Lambda(r(t)) = \lim_{N\to\infty}\frac{p_N(r(t)|H_A)}{p_N(r(t)|H_B)}}$$

$$\boldsymbol{\Lambda(r(t)) = e^{-\sum_{i=0}^{N}\frac{(r(t)-s(t|H_A))^2-(r(t)-s(t|H_B))^2}{N_0}}} \tag{6}$$

The likelihood function is defined as a density function:

$$\lambda(r(t)) = e^{-\sum_{i=0}^{N}\frac{(r(t)-s(t|H_A))^2}{N_0}} \tag{7}$$

Appendix B

BPSK Development

The hypothesis is implicit in the selected model. In the case MPSK signals we have:

$$s(t; H = M) = \sum_{k=0}^{K-1} \sqrt{E}\, e^{j\theta_k} e^{j\theta_o}\, p(t - kT) \quad for\ \ 0 \le t \le NT \tag{8}$$

For unknown types, the phase of the symbol θ_k belongs to a set of M elements:

$$\theta_k = \left\{\frac{2\pi m}{M}\right\}^{m=1,..,M} \tag{9}$$

The pulse is defined as an ideal square pulse with period T. The likelihood can be simplified to:

$$\lambda\big(r(t)\big) = \prod_{k=0}^{K-1} E_{b_k}\left\{e^{\sqrt{\gamma} Re\{b_k e^{-j\theta_c} r_k(\epsilon)\} - \gamma}\right\} \tag{10}$$

where γ is the SNR at the output of the decorrelator and $r_{ki}(\epsilon)$ is the decorrelator output. After some math, the likelihood expression becomes:

$$\lambda(r(t)) = \frac{1}{2} e^{-\gamma} \left(cosh\big(\sqrt{\gamma} Re\{e^{-j\theta_c} r_k(\epsilon)\}\big)\right) \tag{11}$$

BPSK Spread Spectrum Development

This development assumes uniformly distributed coefficients.

$$E_{b_k}\left\{e^{\sqrt{\gamma} Re\{b_k e^{-j\theta_c} \sum_{i=0}^{L-1} c_i^{*}\, r_{ki}(\epsilon)\} - \gamma}\right\} = \int_0^{2\pi} e^{\sqrt{\gamma} Re\{b_k e^{-j\theta_c} \sum_{i=0}^{L-1} c_i^{*}\, r_{ki}(\epsilon)\} - \gamma}\, p(b_k) db_k$$

$$p(\theta_k) = \frac{1}{2} \sum_{m=0}^{1} \delta\left(b_k - e^{j\frac{2\pi m}{M}}\right) \tag{12}$$

For $c_i \in \mathbb{R}$

$$\lambda\big(r(t)\big) = \prod_{k=0}^{K-1}\prod_{i=1}^{L} e^{-\gamma}\cosh\!\left(\sqrt{\gamma}Re\{e^{-j\theta_c} r_{ki}(\epsilon)\}\right) \tag{13}$$

The correlation term $r_{ki}(\epsilon)$ can be substituted by $r_k(\epsilon)$. This means that for this simple case the likelihood of BPSK and BPSK spread spectrum are the same. The ratio test will not be able to differentiate between BPSK and spread spectrum under the assumption of statistically independent symbols.

Appendix C

BPSK Spread Spectrum Development with Weights

The equations were derived using *Mathematica*. This is the likelihood when using a 2x2 Hadamard matrix.

```
In[337]:= (* CDMA *)
Clear[b, c, r]
M = 2;
H = Array[Subscript[c, #1 - 1, #2 - 1] / Sqrt[M] &, {M, M}];
B = Array[Subscript[b, # - 1, K] &, {M, 1}];
R = Array[Subscript[r, K, #1 - 1] &, {M, 1}];
X = Rᵀ.H.B
Y = (H.B)
```

Out[342]= $\left\{\left\{b_{0,K}\left(\frac{c_{0,0}\, r_{K,0}}{\sqrt{2}}+\frac{c_{1,0}\, r_{K,1}}{\sqrt{2}}\right)+b_{1,K}\left(\frac{c_{0,1}\, r_{K,0}}{\sqrt{2}}+\frac{c_{1,1}\, r_{K,1}}{\sqrt{2}}\right)\right\}\right\}$

Out[343]= $\left\{\left\{\frac{b_{0,K}\, c_{0,0}}{\sqrt{2}}+\frac{b_{1,K}\, c_{0,1}}{\sqrt{2}}\right\},\left\{\frac{b_{0,K}\, c_{1,0}}{\sqrt{2}}+\frac{b_{1,K}\, c_{1,1}}{\sqrt{2}}\right\}\right\}$

```
In[344]:= (* WEIGHTS *)
TOT = 0;
For[ i = 0, i < M, i++, For[ j = 0, j < M, j++,
  For[ k = 0, k < M, k++, For[l = 0, l < M, l++,
    If[ i ≠ j && k ≠ l,
     TOT += Subscript[c, i, k] * Subscript[c, i, l] * Subscript[c, j, k] * Subscript[c, j, l];
    ]
   ]]
 ]]
FML = Expand[ ((M^2 - M) M + TOT) / 4]
```

Out[346]= $1 + c_{0,0}\, c_{0,1}\, c_{1,0}\, c_{1,1}$

```
In[347]:= (* ENERGY *)
ET = (Yᵀ.Y / 2);
ET = Expand[ET];
ET = ET /. v_^2 :> 1;
ET
```

Out[350]= $\left\{\left\{1+\frac{1}{2}\, b_{0,K}\, b_{1,K}\, c_{0,0}\, c_{0,1}+\frac{1}{2}\, b_{0,K}\, b_{1,K}\, c_{1,0}\, c_{1,1}\right\}\right\}$

```
In[351]:= (* CONDITIONAL LIKELIHOOD BPSK-SS *)
CL = Exp[Sqrt[2 γ] * (X - FML) - γ ET] [[1, 1]];
Z = Join[Sqrt[M] * Flatten[H], Flatten[B]];
For[ i = 1, i ≤ Length[Z], i++,
 rl = { Z[[i]] → z };
 CL = CL /. rl;
 CL = 1 / 2 * Sum[CL, {z, {-1, 1}}];
]
Expand[CL]
```

Out[354]= $\frac{1}{4} e^{-2\sqrt{2}\sqrt{\gamma}} + \frac{1}{8} e^{-\gamma-2\sqrt{\gamma}\, r_{K,0}} + \frac{1}{8} e^{-\gamma+2\sqrt{\gamma}\, r_{K,0}} + \frac{1}{8} e^{-\gamma-2\sqrt{\gamma}\, r_{K,1}} + \frac{1}{8} e^{-\gamma+2\sqrt{\gamma}\, r_{K,1}} + \frac{1}{16} e^{-2\gamma+\sqrt{2}\sqrt{\gamma}\left(-2-\sqrt{2}\, r_{K,0}-\sqrt{2}\, r_{K,1}\right)} + \frac{1}{16} e^{-2\gamma+\sqrt{2}\sqrt{\gamma}\left(-2+\sqrt{2}\, r_{K,0}-\sqrt{2}\, r_{K,1}\right)} + \frac{1}{16} e^{-2\gamma+\sqrt{2}\sqrt{\gamma}\left(-2-\sqrt{2}\, r_{K,0}+\sqrt{2}\, r_{K,1}\right)} + \frac{1}{16} e^{-2\gamma+\sqrt{2}\sqrt{\gamma}\left(-2+\sqrt{2}\, r_{K,0}+\sqrt{2}\, r_{K,1}\right)}$

BPSK Spread Spectrum Development with Weights

The equations were derived using *Mathematica*. This is the likelihood when using a4x4 Hadamard matrix.

```
In[319]:= (* CDMA *)
          Clear[b, c, r]
          M = 4;
          H = Array[Subscript[c, #1 - 1, #2 - 1] / Sqrt[M] &, {M, M}];
          B = Array[Subscript[b, # - 1, K] &, {M, 1}];
          R = Array[Subscript[r, K, #1 - 1] &, {M, 1}];
          X = R^T.H.B
          Y = (H.B)
```

Out[324]=
$$\left\{\left\{b_{0,K}\left(\frac{1}{2}c_{0,0}r_{K,0}+\frac{1}{2}c_{1,0}r_{K,1}+\frac{1}{2}c_{2,0}r_{K,2}+\frac{1}{2}c_{3,0}r_{K,3}\right)+b_{1,K}\left(\frac{1}{2}c_{0,1}r_{K,0}+\frac{1}{2}c_{1,1}r_{K,1}+\frac{1}{2}c_{2,1}r_{K,2}+\frac{1}{2}c_{3,1}r_{K,3}\right)+b_{2,K}\left(\frac{1}{2}c_{0,2}r_{K,0}+\frac{1}{2}c_{1,2}r_{K,1}+\frac{1}{2}c_{2,2}r_{K,2}+\frac{1}{2}c_{3,2}r_{K,3}\right)+b_{3,K}\left(\frac{1}{2}c_{0,3}r_{K,0}+\frac{1}{2}c_{1,3}r_{K,1}+\frac{1}{2}c_{2,3}r_{K,2}+\frac{1}{2}c_{3,3}r_{K,3}\right)\right\}\right\}$$

Out[325]=
$$\left\{\left\{\frac{1}{2}b_{0,K}c_{0,0}+\frac{1}{2}b_{1,K}c_{0,1}+\frac{1}{2}b_{2,K}c_{0,2}+\frac{1}{2}b_{3,K}c_{0,3}\right\},\left\{\frac{1}{2}b_{0,K}c_{1,0}+\frac{1}{2}b_{1,K}c_{1,1}+\frac{1}{2}b_{2,K}c_{1,2}+\frac{1}{2}b_{3,K}c_{1,3}\right\},\left\{\frac{1}{2}b_{0,K}c_{2,0}+\frac{1}{2}b_{1,K}c_{2,1}+\frac{1}{2}b_{2,K}c_{2,2}+\frac{1}{2}b_{3,K}c_{2,3}\right\},\left\{\frac{1}{2}b_{0,K}c_{3,0}+\frac{1}{2}b_{1,K}c_{3,1}+\frac{1}{2}b_{2,K}c_{3,2}+\frac{1}{2}b_{3,K}c_{3,3}\right\}\right\}$$

```
In[326]:= (* WEIGHTS *)
          TOT = 0;
          For[ i = 0, i < M, i++, For[ j = 0, j < M, j++,
            For[ k = 0, k < M, k++, For[l = 0, l < M, l++,
              If[ i ≠ j && k ≠ l,
               TOT += Subscript[c, i, k] * Subscript[c, i, l] * Subscript[c, j, k] * Subscript[c, j, l];
              ]
             ]]
           ]]
          FML = Expand[((M^2 - M) M + TOT) / 4]
```

Out[328]=
$$12+c_{0,0}c_{0,1}c_{1,0}c_{1,1}+c_{0,0}c_{0,2}c_{1,0}c_{1,2}+c_{0,1}c_{0,2}c_{1,1}c_{1,2}+c_{0,0}c_{0,3}c_{1,0}c_{1,3}+c_{0,1}c_{0,3}c_{1,1}c_{1,3}+c_{0,2}c_{0,3}c_{1,2}c_{1,3}+c_{0,0}c_{0,1}c_{2,0}c_{2,1}+c_{1,0}c_{1,1}c_{2,0}c_{2,1}+c_{0,0}c_{0,2}c_{2,0}c_{2,2}+c_{1,0}c_{1,2}c_{2,0}c_{2,2}+c_{0,1}c_{0,2}c_{2,1}c_{2,2}+c_{1,1}c_{1,2}c_{2,1}c_{2,2}+c_{0,0}c_{0,3}c_{2,0}c_{2,3}+c_{1,0}c_{1,3}c_{2,0}c_{2,3}+c_{0,1}c_{0,3}c_{2,1}c_{2,3}+c_{1,1}c_{1,3}c_{2,1}c_{2,3}+c_{0,2}c_{0,3}c_{2,2}c_{2,3}+c_{1,2}c_{1,3}c_{2,2}c_{2,3}+c_{0,0}c_{0,1}c_{3,0}c_{3,1}+c_{1,0}c_{1,1}c_{3,0}c_{3,1}+c_{2,0}c_{2,1}c_{3,0}c_{3,1}+c_{0,0}c_{0,2}c_{3,0}c_{3,2}+c_{1,0}c_{1,2}c_{3,0}c_{3,2}+c_{2,0}c_{2,2}c_{3,0}c_{3,2}+c_{0,1}c_{0,2}c_{3,1}c_{3,2}+c_{1,1}c_{1,2}c_{3,1}c_{3,2}+c_{2,1}c_{2,2}c_{3,1}c_{3,2}+c_{0,0}c_{0,3}c_{3,0}c_{3,3}+c_{1,0}c_{1,3}c_{3,0}c_{3,3}+c_{2,0}c_{2,3}c_{3,0}c_{3,3}+c_{0,1}c_{0,3}c_{3,1}c_{3,3}+c_{1,1}c_{1,3}c_{3,1}c_{3,3}+c_{2,1}c_{2,3}c_{3,1}c_{3,3}+c_{0,2}c_{0,3}c_{3,2}c_{3,3}+c_{1,2}c_{1,3}c_{3,2}c_{3,3}+c_{2,2}c_{2,3}c_{3,2}c_{3,3}$$

```
In[329]:= (* ENERGY *)
ET = (Y^T.Y / 2);
ET = Expand[ET];
ET = ET /. v_^2 :> 1;
ET
```

Out[332]=

$$\Big\{\Big\{2 + \frac{1}{4} b_{0,K} b_{1,K} c_{0,0} c_{0,1} + \frac{1}{4} b_{0,K} b_{2,K} c_{0,0} c_{0,2} + \frac{1}{4} b_{1,K} b_{2,K} c_{0,1} c_{0,2} + \frac{1}{4} b_{0,K} b_{3,K} c_{0,0} c_{0,3} +$$
$$\frac{1}{4} b_{1,K} b_{3,K} c_{0,1} c_{0,3} + \frac{1}{4} b_{2,K} b_{3,K} c_{0,2} c_{0,3} + \frac{1}{4} b_{0,K} b_{1,K} c_{1,0} c_{1,1} + \frac{1}{4} b_{0,K} b_{2,K} c_{1,0} c_{1,2} +$$
$$\frac{1}{4} b_{1,K} b_{2,K} c_{1,1} c_{1,2} + \frac{1}{4} b_{0,K} b_{3,K} c_{1,0} c_{1,3} + \frac{1}{4} b_{1,K} b_{3,K} c_{1,1} c_{1,3} + \frac{1}{4} b_{2,K} b_{3,K} c_{1,2} c_{1,3} +$$
$$\frac{1}{4} b_{0,K} b_{1,K} c_{2,0} c_{2,1} + \frac{1}{4} b_{0,K} b_{2,K} c_{2,0} c_{2,2} + \frac{1}{4} b_{1,K} b_{2,K} c_{2,1} c_{2,2} + \frac{1}{4} b_{0,K} b_{3,K} c_{2,0} c_{2,3} +$$
$$\frac{1}{4} b_{1,K} b_{3,K} c_{2,1} c_{2,3} + \frac{1}{4} b_{2,K} b_{3,K} c_{2,2} c_{2,3} + \frac{1}{4} b_{0,K} b_{1,K} c_{3,0} c_{3,1} + \frac{1}{4} b_{0,K} b_{2,K} c_{3,0} c_{3,2} +$$
$$\frac{1}{4} b_{1,K} b_{2,K} c_{3,1} c_{3,2} + \frac{1}{4} b_{0,K} b_{3,K} c_{3,0} c_{3,3} + \frac{1}{4} b_{1,K} b_{3,K} c_{3,1} c_{3,3} + \frac{1}{4} b_{2,K} b_{3,K} c_{3,2} c_{3,3}\Big\}\Big\}$$

```
In[333]:= (* CONDITIONAL LIKELIHOOD BPSK-SS *)
CL = Exp[Sqrt[2 γ] * (X - FML) - γ ET][[1, 1]];
Z = Join[Sqrt[M] * Flatten[H], Flatten[B]];
For[ i = 1, i ≤ Length[Z], i++,
 rl = { Z[[i]] → z };
 CL = CL /. rl;
 CL = 1 / 2 * Sum[CL, {z, {-1, 1}}];
]
Expand[CL]
```

Out[336]=

A very large output was generated. Here is a sample of it:

$$\frac{3\, e^{-16\sqrt{2}\sqrt{\gamma}}}{4096} + \frac{3}{512} e^{-24\sqrt{2}\sqrt{\gamma}} + \frac{9\, e^{-16\sqrt{2}\sqrt{\gamma}}}{2048} + \frac{9\, e^{-8\sqrt{2}\sqrt{\gamma}}}{1024} + \ll 1221 \gg +$$
$$\frac{e^{-\frac{11\gamma}{2} - \sqrt{2}\sqrt{\gamma}\,(\ldots)}}{16384} + \frac{e^{-\frac{11\gamma}{2} - \sqrt{2}\sqrt{\gamma}\,(\ldots)}}{16384} + \frac{e^{-8\gamma - \sqrt{2}\sqrt{\gamma}\,(\ldots)}}{65536}$$

Show Less | Show More | Show Full Output | Set Size Limit...

The likelihood function has 1228 exponential terms. The terms are combinations of the vectors components of the spread spectrum signal.

List of Acronyms

ALRT	Average LRT
BPSK	Binary Shift Keying
BPSK-SS	BPSK Spread Spectrum or CDMA
D_{KL}	Kullback-Leibler Information Divergence
GLRT	Generalized LRT
HLRT	Hybrid LRT
LRT	Likelihood Ratio Test
MIMO	Multiple Input Multiple Output System
MPSK	M-ary Shift Keying
No	Normalized Noise Power
Pdf	Probability Density Function
PEP	Pairwise Error Probability
QAM	Quadrature Amplitude Modulation
ROC	Receiver Operating Characteristic
STBC	Space-Time Block Codes
TSC	Total Squared Correlation
θ	Golden Ratio
$\mathbb{R}$	Real Numbers

$\mathbb{Q}$	Rational Numbers
$\mathbb{Z}$	Complex Numbers
$\mathbb{Z}[i]$	Gaussian Integers
$\mathcal{N}(\mathcal{N}, \sigma^2)$	Gaussian Distribution with mean μ and variance σ^2